

**Universidad Carlos III de Madrid**

**Escuela Politécnica Superior**



**Ingeniería Informática (2º Ciclo)**

**Proyecto Fin de Carrera**

**HERRAMIENTA PARA LA EXTRACCIÓN, CONSULTA Y  
REPRESENTACIÓN DE EVENTOS EN PUBLICACIONES  
CIENTÍFICAS MEDIANTE LAS APIs DE GOOGLE.**

**Autora:** Sonia Morcillo Muñoz

**Tutora:** María Teresa Vicente Díez

Diciembre, 2009



## **Agradecimientos**

Tengo que agradecer el haber podido llevar a cabo el desarrollo de este proyecto a mi amiga, a la par que antigua compañera y jefa de equipo, Mayte. Gracias a ella he tenido el ánimo y la fuerza suficiente para sacarlo adelante. Gracias.

También tengo que dar las gracias a algunos compañeros y amigos por el apoyo que he recibido en algunos momentos difíciles que se han dado a lo largo del proyecto.

A todos, gracias.

## Índice de contenidos

<b>1</b>	<b>INTRODUCCIÓN .....</b>	<b>7</b>
1.1	MOTIVACIÓN .....	7
1.2	OBJETIVOS.....	8
1.3	ESTRUCTURA DEL DOCUMENTO.....	9
<b>2</b>	<b>ESTADO DEL ARTE.....</b>	<b>13</b>
2.1	INTRODUCCIÓN A LOS <i>CALL FOR PAPERS</i> (CFP) .....	13
2.1.1	<i>CONFERENCIAS ACADÉMICAS Y CFP</i> .....	13
2.1.2	<i>ORGANIZACIÓN DE UNA CONFERENCIA ACADÉMICA Y USO DE CALL FOR PAPERS (CFP)</i> .....	14
2.1.3	<i>ALGUNOS EJEMPLOS DE PÁGINAS WEB QUE RECOPILAN INFORMACIÓN DE CONFERENCIAS CIENTÍFICAS</i> .....	15
2.2	EXTRACCIÓN DE INFORMACIÓN .....	22
2.2.1	<i>CONCEPTOS RELACIONADOS CON LA EXTRACCIÓN DE INFORMACIÓN</i> .....	22
2.2.2	<i>WEB 2.0 y WEB 3.0</i> .....	31
2.2.3	<i>MICROFORMATOS</i> .....	32
2.2.4	<i>EJEMPLO DE APLICACIÓN DEL MICROFORMATO HCALENDAR</i> .....	38
2.3	APIs DE GOOGLE .....	41
2.3.1	<i>API DE DATOS DE GOOGLE (GOOGLE DATA API)</i> .....	41
2.3.2	<i>API DE DATOS DE GOOGLE CALENDAR</i> .....	45
2.3.3	<i>API AJAX DE GOOGLE</i> .....	58
2.3.4	<i>API DE GOOGLE MAPS</i> .....	61
2.3.5	<i>API AJAX DE GOOGLE SEARCH</i> .....	68
2.4	HERRAMIENTAS Y TECNOLOGÍAS QUE SE UTILIZAN EN EL SISTEMA .....	71
2.4.1	<i>MICROSOFT VISUAL STUDIO</i> .....	71
2.4.2	<i>EXPRESIONES REGULARES</i> .....	72
2.4.3	<i>REGULATOR 2.03</i> .....	74
2.4.4	<i>HTML</i> .....	74
2.4.5	<i>XML</i> .....	75
2.4.6	<i>JAVASCRIPT</i> .....	77
<b>3</b>	<b>ESTUDIO DE LA VIABILIDAD Y PLANTEAMIENTO DEL PROYECTO .....</b>	<b>79</b>
3.1	DESCRIPCIÓN GENERAL DEL SISTEMA .....	79
3.2	DESCRIPCIÓN DEL ENTORNO, USUARIOS Y USOS. ....	80
3.3	ALCANCE DEL PROYECTO.....	81
3.3.1	<i>REQUISITOS DE USUARIO</i> .....	83
3.3.2	<i>REQUISITOS DEL SISTEMA</i> .....	84
3.4	ANÁLISIS DE VIABILIDAD .....	88
3.4.1	<i>SOFTWARE</i> .....	88

3.4.2	<i>HARDWARE</i> .....	89
3.5	PRUEBAS DE VALIDACIÓN .....	90
3.6	METODOLOGÍA DEL DESARROLLO .....	99
<b>4</b>	<b>ANÁLISIS Y DISEÑO DEL SISTEMA .....</b>	<b>102</b>
4.1	ESPECIFICACIÓN DE REQUISITOS.....	104
4.1.1	<i>REQUISITOS DE USUARIO</i> .....	104
4.1.2	<i>REQUISITOS DE SISTEMA</i> .....	108
4.2	DIAGRAMA DE CASOS DE USO .....	148
4.3	DIAGRAMAS DE SECUENCIA.....	157
4.4	DIAGRAMA DE CLASES .....	169
4.4.1	<i>DESCRIPCIÓN DE LAS CLASES DEL SISTEMA</i> .....	170
4.4.2	<i>CLASE DEFINICIONES</i> .....	171
4.4.3	<i>CLASE EXTRACCIONINFORMACION</i> .....	172
4.4.4	<i>CLASE CALENDAR</i> .....	174
4.4.5	<i>CLASE CONFIGURACION</i> .....	177
4.4.6	<i>CLASE FRMEVENTO</i> .....	177
4.4.7	<i>CLASE FRMPRINCIPAL</i> .....	178
4.4.8	<i>CLASE GMAPS</i> .....	178
4.4.9	<i>CLASE TRAZALOG</i> .....	179
4.4.10	<i>CLASE TRIPLEDES</i> .....	179
<b>5</b>	<b>IMPLEMENTACIÓN DEL SISTEMA .....</b>	<b>181</b>
5.1	PESTAÑA DE EXTRACCIÓN DE INFORMACIÓN.....	181
5.2	PESTAÑA DE GESTIÓN DE EVENTOS DE CALENDARIO DE GOOGLE CALENDAR 186	
5.3	PESTAÑA DE BÚSQUEDA E INSERCIÓN MANUAL DE EVENTOS .....	188
5.4	PESTAÑA DE CONFIGURACIÓN .....	189
<b>6</b>	<b>PRUEBAS.....</b>	<b>191</b>
6.1	EJECUCIÓN DE PRUEBAS .....	191
6.2	ACCIONES CORRECTIVAS.....	192
<b>7</b>	<b>CONCLUSIONES .....</b>	<b>193</b>
<b>8</b>	<b>LÍNEAS FUTURAS.....</b>	<b>195</b>
<b>9</b>	<b>LISTA DE ACRÓNIMOS.....</b>	<b>197</b>
<b>10</b>	<b>BIBLIOGRAFÍA .....</b>	<b>199</b>
10.1	MATERIAL REFERENCIADO .....	199
10.2	MATERIAL DE APOYO.....	200
<b>11</b>	<b>ANEXOS.....</b>	<b>202</b>

ANEXO 1 – MANUAL DE USUARIO.....202

ANEXO 2 – MANUAL DE INSTALACIÓN .....221

ANEXO 3 – ÍNDICE DE FIGURAS .....228

ANEXO 4 – ÍNDICE DE TABLAS .....231

# **1 INTRODUCCIÓN**

## **1.1 MOTIVACIÓN**

La extracción, organización y seguimiento de eventos relacionados con las publicaciones en conferencias científicas es una tarea complicada, por su heterogeneidad en las fuentes de consulta, y por su dispersión, puesto que cada conferencia suele tener su propia página Web donde publica sus eventos y sólo algunas Webs recopilan esta información, de forma heterogénea. Es por ello que surge la necesidad de encontrar una solución que ayude a los investigadores a gestionar su calendario de publicaciones.

A través de este proyecto fin de carrera se pretende contribuir a la organización de las necesidades de gestión y almacenamiento de información proveniente del ámbito de las publicaciones científicas, creando un sistema que facilite la extracción y gestión de los eventos que dichas publicaciones llevan asociadas.

Para conseguir esto, es necesario un estudio previo de las posibilidades de extracción automática de información de páginas Web de conferencias científicas. La idea es obtener los eventos más relevantes publicados en la Web de forma automatizada, realizando un análisis de las fuentes de información en el ámbito de las conferencias científicas, y de los hitos relacionados que resultan más destacables.

En cuanto a la forma de almacenamiento y gestión de los eventos dentro del sistema, se propone la utilización de las funcionalidades del API de Datos de Google [1], con servicios como Google Calendar para ubicar y representar temporalmente los eventos. Se plantea asimismo el uso de las APIs de AJAX de Google [2] como Google Maps para mostrar geográficamente las ubicaciones asociadas en un mapa y Google Search para facilitar la integración de un buscador de información Web.

La realización de esta parte del proyecto implica por tanto un estudio profundo de las funcionalidades de dichas APIs.

## 1.2 OBJETIVOS

Los principales objetivos que se abordan en este proyecto, son los siguientes:

- Investigación sobre las técnicas de extracción de información automática de contenidos Web semiestructurados.
- Familiarización con la información y la nomenclatura manejada en las páginas Web relacionadas con eventos de conferencias científicas.
- Estudio y aplicación del API de Datos de Google Calendar y del API AJAX de Google con vistas a resolver el problema de la gestión de información de conferencias.
- Aplicación de lo analizado en los puntos anteriores mediante la creación de una herramienta, que permita la extracción, consulta y representación de eventos frecuentemente manejados en las llamadas a la publicación de conferencias científicas, en una aplicación de escritorio. La herramienta construida deberá suponer un aporte en las siguiente líneas:
  1. Automatizar el proceso de obtención y almacenamiento de información sobre determinadas páginas que contienen eventos relacionados con las publicaciones en conferencias científicas.
  2. Insertar los eventos obtenidos en el proceso anterior, en un calendario Web online utilizando el API de Datos de Google Calendar.
  3. Obtener los eventos guardados en el calendario de Google Calendar y mostrarlos en el sistema, facilitando filtros de búsqueda a la hora de su consulta.



4. Gestionar estos eventos que se encuentran en el calendario de Google Calendar desde la herramienta, con la posibilidad de modificar sus datos, eliminarlos e insertar manualmente nuevos eventos.
  5. Ubicar geográficamente los lugares de celebración de los eventos, utilizando para ello mapas de Google Maps.
  6. Facilitar un buscador de información dentro de la herramienta, utilizando para ello las funcionalidades del API de Google Search. De esta manera, se permitirá si se desea, realizar búsquedas de nuevos eventos en toda la Web de forma rápida, con la finalidad de poder insertarlos manualmente y almacenarlos a través de la aplicación.
- Generación de una documentación completa del ciclo de desarrollo de la herramienta, así como una guía de su funcionamiento e instalación que facilite su uso por aquéllos miembros de la comunidad científica interesados en la propuesta.

### **1.3 ESTRUCTURA DEL DOCUMENTO**

A continuación se describe en líneas generales, el contenido de cada capítulo incluido en esta memoria:

- Capítulo 1 – Introducción

En este capítulo se comenta la motivación para la realización del proyecto, los objetivos que se pretenden llevar a cabo, y el punto actual de estructura del documento.

- El Capítulo 2 - Estado del Arte

En este capítulo se especifican algunos aspectos teóricos relacionados con el proyecto:

- Un apartado de introducción a los *Call For Papers* y su estrecha relación con las publicaciones de conferencias científicas.
- Un apartado relacionado con el tema de la extracción automática de información. En este punto también se habla de los conceptos de Web 2.0 y Web 3.0., de los microformatos y su importancia dentro de la Web semántica.
- Un apartado que contiene la descripción y contenido de información relacionada con el API de Google, concretando la información concerniente a las APIs que se utilizan para el desarrollo de este proyecto: API de Google Calendar, API de Google Maps y API AJAX de Google Search.
- Un último apartado donde se incluyen conceptos y documentación relacionados con las herramientas y la tecnología asociada al desarrollo de este proyecto.

### Capítulo 3 – Estudio de la viabilidad y planteamiento del proyecto

En este capítulo se realiza la descripción general del sistema, la descripción del entorno, y el alcance del proyecto. Se incluye también el análisis de viabilidad, tanto de software como de hardware, la propuesta de pruebas de validación, y se especifica la metodología de desarrollo a utilizar para la realización del proyecto.

### Capítulo 4 – Análisis y Diseño del Sistema

En este capítulo se especifica el catálogo de requisitos completo, los diagramas de casos de uso y de secuencia, y el diagrama de clases.

### Capítulo 5 – Implementación del Sistema

En este capítulo se especifican aspectos relacionados con la implementación llevada a cabo para el desarrollo del proyecto.

## Capítulo 6 – Pruebas

En este capítulo se documentan los resultados de la ejecución de las pruebas de validación realizadas en el sistema, y las acciones correctivas, si procede.

## Capítulo 7 – Conclusiones

En este capítulo se comentan las conclusiones obtenidas de lo aprendido tras la finalización del proyecto.

## Capítulo 8 – Líneas Futuras

En este capítulo se comentan las posibles ampliaciones y mejoras que se pueden realizar en un futuro sobre la aplicación.

## Capítulo 9 – Lista de Acrónimos

En este capítulo se incluyen todos los acrónimos que se han mencionado en algún momento en esta memoria, con su significado.

## Capítulo 10 – Bibliografía

En este capítulo se especifica la bibliografía utilizada para el desarrollo del proyecto. Consta de una parte de referencias obtenidas de la Web (documentación *on-line*), y otra parte de referencias bibliográficas.

## Capítulo 11 – Anexos

Este capítulo incluye el manual de usuario y el manual de instalación de la aplicación, así como los índices de ilustraciones y tablas incluidos en la memoria del proyecto.

## **2 ESTADO DEL ARTE**

### **2.1 INTRODUCCIÓN A LOS *CALL FOR PAPERS* (CFP)**

#### **2.1.1 CONFERENCIAS ACADÉMICAS Y CFP**

Una conferencia académica es una forma de divulgación en la que los investigadores (no siempre académicos) presentan y discuten sus trabajos. Junto con publicaciones científicas o académicas, las conferencias proporcionan un importante canal para el intercambio de información entre investigadores.

Generalmente, el trabajo se expone en forma de presentaciones concisas que duran entre 10 y 30 minutos, y que normalmente incluyen un debate posterior. El trabajo puede ser organizado en forma de artículos académicos (comúnmente denominados “*papers*”), que son publicados en las actas de la conferencia. A menudo hay uno o más ponentes destacados (normalmente expertos con prestigio), que realizan una ponencia principal y que comúnmente se anuncia antes de la conferencia. Las mesas redondas y los talleres también pueden ser parte de una conferencia académica, particularmente si la conferencia está relacionada con las artes interpretativas.

A los futuros ponentes normalmente se les pide que envíen un pequeño resumen de sus presentaciones, siendo revisados antes de que la presentación sea aceptada para el encuentro. Algunas disciplinas requieren que los asistentes que deseen presentar un trabajo en la conferencia envíen un artículo de 6-15 páginas, que será revisado por los miembros del comité del programa o por revisores elegidos por ellos.

En algunas temáticas, como inglés u otros idiomas, es común que los ponentes hagan la presentación a partir de un guión preparado previamente. En otras temáticas como las ciencias, los ponentes suelen basar sus exposiciones en presentaciones visuales que muestran figuras clave y resultados de su investigación.

Un encuentro largo normalmente se llama conferencia, mientras que uno más corto se suele denominar *workshop* o taller. Estos encuentros pueden ser *single track* o *multiple track*. En el primer caso, sólo se celebra una sesión a la vez, mientras que un encuentro *multiple track*, se celebran varias sesiones en paralelo y varios ponentes presentan su trabajo en habitaciones separadas de manera simultánea.

Dependiendo del tema de la conferencia, se pueden ofrecer también actividades sociales (*social events*) o de entretenimiento. Si la conferencia es bastante larga, las editoriales académicas pueden también convocar exposiciones y ofrecer sus libros. En las conferencias más largas también tienen lugar encuentros de negocio para sociedades de expertos o grupos interesados.

En las conferencias académicas se consideran tres categorías:

- Las conferencias temáticas, son pequeñas conferencias organizadas alrededor de un tema particular.
- La conferencia general, con un enfoque más amplio, con sesiones de una amplia variedad de temas. Estas conferencias se organizan a menudo por sociedades de eruditos internacionales, nacionales o regionales, y se celebran con periodicidad regular.
- La conferencia profesional, conferencias más largas no limitadas al ámbito académico, pero con temas relacionados con dicho entorno.

### **2.1.2 ORGANIZACIÓN DE UNA CONFERENCIA ACADÉMICA Y USO DE CALL FOR PAPERS (CFP)**

Las conferencias son organizadas normalmente por sociedades científicas o por un grupo de investigadores con un interés común. Los encuentros más largos suelen ser gestionados por una sociedad científica a través de un Organizador de Conferencias Profesionales o PCO.

El encuentro se anuncia a través de un *Call For Papers* (CFP) o un *Call For Abstracts* (CFA), que enumera los temas de dicho encuentro e indica a los futuros ponentes cómo enviar sus resúmenes (*abstracts*) o artículos (*papers*). Los envíos tienen lugar de manera online usando un servicio de entregas.

Por tanto, se podría definir más ampliamente el CFP como un método usado en el entorno académico y en otros contextos para atraer el interés de la comunidad en la publicación de capítulos de libros, artículos o presentaciones dentro de conferencias. Un CFP normalmente se envía a los grupos interesados describiendo el tema general, el motivo para el CFP, formalidades como qué clase de resumen (*abstract*) tiene que ser enviado, a quién y un *deadline* o fecha tope de envío.


Un CFP suele distribuirse usando una lista de correo o mediante servicios online especializados.

### **2.1.3 ALGUNOS EJEMPLOS DE PÁGINAS WEB QUE RECOPILAN INFORMACIÓN DE CONFERENCIAS CIENTÍFICAS**

Actualmente existen varias páginas Web que recopilan información sobre los eventos de conferencias científicas. Las páginas que se han evaluado para el desarrollo de este proyecto contienen CFPs de conferencias relacionadas con las tecnologías de la información y con el entorno de la informática.

A continuación se muestran algunas de ellas:


- **Página de eventos “WikiCFP” [11] :**


A Wiki for Calls For Papers

[Home](#)  
[Login](#)  
[Register](#)  
[Profile](#)  
[Logout](#)  
[Categories](#)  
[Calls For Papers](#)  
[Post a CFP](#)  
[My List](#)  
[Timeline](#)  
[My Archive](#)  
[On iPhone](#)

**What's Special**  
1. customized list  
2. category view  
3. sort by dates  
4. search by title, location or category  
5. rss subscription to lists or categories  
6. interactive timeline  
7. iCalendar links  
8. email reminders  
9. by researchers and for researchers

**Related Resources**  
[ACM CFPs](#)  
[IEEE CFPs](#)

[Visitor locations](#)  
  
[Click to see](#)  
[follow us on twitter](#)


**Calls For Papers on NLP**  
[SHARE](#) [RSS FEED](#)  
(ordered by deadline)

Event	When	Where	Deadline
<a href="#">ICSC 2010</a>	Fourth IEEE International Conference on Semantic Computing Sep 20, 2010 - Sep 22, 2010	Pittsburgh, PA, USA	May 3, 2010
<a href="#">INTERSPEECH 2010</a>	International Conference on Spoken Language Processing Sep 26, 2010 - Sep 30, 2010	Makuhari, Japan	Apr 30, 2010 (Jul 28, 2010)
<a href="#">KONVENS 2010</a> <b>NEW</b>	Conference on Natural Language Processing Sep 6, 2010 - Sep 8, 2010	Saarbrücken, Germany	Apr 23, 2010
<a href="#">COLING 2010</a>	23rd International Conference on Computational Linguistics Aug 23, 2010 - Aug 27, 2010	Beijing, China	Apr 19, 2010
<a href="#">CANLI 2010</a>	Book chapter for the book: Conversational Agents and Natural Language Interaction: Techniques and Effective Practices N/A	N/A	Apr 16, 2010 (Dec 16, 2009)
<a href="#">DLT 2010</a>	14th International Conference on Developments in Language Theory Aug 17, 2010 - Aug 20, 2010	London, Ontario, Canada	Mar 20, 2010
<a href="#">NLG 2010</a>	6th International Natural Language Generation Conference May 7, 2010 - Jul 9, 2010	Trim Castle, Ireland	Mar 14, 2010
<a href="#">CoNLL 2010</a>	Fourteenth Conference on Computational Natural Language Learning Jul 15, 2010 - Jul 16, 2010	Uppsala, Sweden	Mar 8, 2010
<a href="#">NLPCS 2010</a>	The 7th International Workshop on Natural Language Processing and Cognitive Science Jun 8, 2010 - Jun 9, 2010	Funchal, Madeira, Portugal	Mar 8, 2010
<a href="#">topiCS-PRE 2010</a>	Journal Topics in Cognitive Science (topiCS), PRODUCTION OF REFERRING EXPRESSIONS: Bridging the gap between computational and empirical approaches to reference N/A	N/A	Mar 1, 2010 (Feb 1, 2010)
<a href="#">BEA 2010</a> <b>NEW</b>	The 5th Workshop on the Innovative Use of NLP for Building Educational Applications Jun 5, 2010 - Jun 6, 2010	Los Angeles, California	Mar 1, 2010
<a href="#">3DVNL 2010</a>	KES2010 workshop on 3D Visualisation of Natural Language Sep 8, 2010 - Sep 10, 2010	Cardiff University, UK	Mar 1, 2010
<a href="#">SLPAT 2010</a> <b>NEW</b>	First Workshop on Speech and Language Processing for Assistive Technologies Jun 5, 2010 - Jun 6, 2010	Los Angeles, CA	Mar 1, 2010
<a href="#">Louhi 2010</a> <b>NEW</b>	Second Louhi Workshop on Text and Data Mining of Health Documents Jun 5, 2010 - Jun 6, 2010	Los Angeles, CA, USA	Mar 1, 2010
<a href="#">CAAGET 2010</a> <b>NEW</b>	NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text Jun 5, 2010 - Jun 6, 2010	Los Angeles, CA, USA	Mar 1, 2010
<a href="#">CALC 2010</a> <b>NEW</b>	NAACL HLT 2010 Workshop on Computational Approaches to Linguistic Creativity May 5, 2010 - Jun 6, 2010	Los Angeles, CA, USA	Mar 1, 2010
<a href="#">SPMRL 2010</a> <b>NEW</b>	First Workshop on Statistical Parsing of Morphologically Rich Languages Jun 5, 2010 - Jun 6, 2010	Los Angeles, CA, USA	Mar 1, 2010
<a href="#">CL&amp;W 2010</a> <b>NEW</b>	NAACL-HLT 2010 Workshop on Computational Linguistics and Writing: Writing Processes and Authoring Aids Jun 5, 2010 - Jun 6, 2010	Los Angeles, CA, USA	Mar 1, 2010
<a href="#">SemSearch 2010</a> <b>NEW</b>	NAACL HLT 2010 Workshop on Semantic Search Jun 5, 2010 - Jun 6, 2010	Los Angeles, CA	Mar 1, 2010
<a href="#">ACL 2010</a>	The 48th Annual Meeting of the Association for Computational Linguistics Jul 11, 2010 - Jul 16, 2010	Uppsala, Sweden	Feb 15, 2010

[first](#) | [previous](#) | [Page 1](#) | [next](#) | [last](#)

Total of 338 CFPs in 17 pages

[About Us](#) | [Contact Us](#) | [Privacy Policy](#) | [Terms and Conditions](#)  
This wiki is licensed under a [Creative Commons Attribution-Share Alike 3.0 License](#).



**Figura 1.** Página de eventos de conferencias científicas del tipo “WikiCFP”

En esta página la información sobre los eventos está semiestructurada. Los datos se encuentran insertados en tablas siguiendo un patrón más o menos común. Aunque en algún caso, como las fechas de *deadline*, hay datos de difícil interpretación. Por ejemplo en el caso de las fechas: “Apr 16, 2010 (Dec 16, 2009)”. En ningún sitio se indica de qué forma interpretar las dos fechas.



- **Página de eventos “ACLWiki” [12]:**

The screenshot shows the ACLWiki website's 'Conferences in 2009' page. The page has a header with navigation tabs (article, discussion, view source, history) and a user login link. A sidebar on the left contains navigation links (Main Page, Community portal, Current events, Recent changes, Random page, Help, Donations) and a search box. The main content area lists 14 conferences in chronological order, each with a brief description, location, and dates. The conferences listed are IWCS-8, CICLing 2009 + Lexicom Americas 2009, EACL 2009, PALC 2009, YLMP 2009, NODALIDA-2009, ICWSM-2009, NAACL-HLT 2009, FSMNLP 2009, ACL-IJCNLP 2009, SEPLN 2009, RANLP-2009, CEE-SECR 2009, and LTC 2009. The page includes a footer with a timestamp, access statistics, and links to privacy policy, about ACLWiki, and disclaimers.

article discussion view source history

Log in / create account

## Conferences in 2009

For more information, see [Conferences and workshops](#). Please add conferences in chronological order, in the following format:

**URL with Short Name as a heading**

- Long Name or Description
- Location
- Dates (submission date)

**Contents [hide]**

- IWCS-8
- CICLing 2009 + Lexicom Americas 2009
- EACL 2009
- PALC 2009
- YLMP 2009
- NODALIDA-2009
- ICWSM-2009
- NAACL-HLT 2009
- FSMNLP 2009
- ACL-IJCNLP 2009
- SEPLN 2009
- RANLP-2009
- CEE-SECR 2009
- LTC 2009

**IWCS-8**

- 8th International Conference on Computational Semantics
- Tilburg, the Netherlands
- Jan 7-9, 2009

**CICLing 2009 + Lexicom Americas 2009**

- 10th International Conference on Intelligent Text Processing and Computational Linguistics, preceded by the Lexicom Americas workshop
- Mexico
- CICLing: March 1-7, 2009 (submission deadline: Nov 3 registration of abstracts, Nov 10 submission of full text). Lexicom: Feb 24-28

**EACL 2009**

- 12th Conference of the European Chapter of the ACL
- Athens, Greece
- March 30-April 3, 2009 (submission deadline: Oct 10, 2008)

**PALC 2009**

- 7th Conference on Practical Applications in Language and Computers
- Lodz, Poland
- April 6-8, 2009 (submission deadline: Jan 31, 2009)

**YLMP 2009**

- 1st Young Linguists' Meeting in Poznan
- student conference focused on interdisciplinary linguistics
- Poznań, Poland
- April 24-26, 2009 (submission deadline: Feb 15, 2009)

**NODALIDA-2009**

- 17th Nordic Conference of Computational Linguistics
- Odense, Denmark
- May 14-16, 2009

**ICWSM-2009**

- International Conference on Weblogs and Social Media 2009
- San Jose, California, U.S.A.
- May 17-20, 2009 (submission deadline: main conference: Jan 21, 2009; workshop Mar 1, 2009)

**NAACL-HLT 2009**

- Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL HLT) 2009
- Boulder, Colorado
- May 31-June 5, 2009

**FSMNLP 2009**

- Finite-State Methods and Natural Language Processing - FSMNLP 2009
- Pretoria, South Africa
- July 21-24, 2009

**ACL-IJCNLP 2009**

- 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing
- Singapore
- August 2-7, 2009 (submission deadline: Feb 22, 2009)

**SEPLN 2009**

- 25th Annual Congress of the [SEPLN](#) Sociedad Española para el Procesamiento del Lenguaje Natural (Spanish Association for Natural Language Processing)
- Donostia - San Sebastián, University of The Basque Country
- September 8-10, 2009 (submission deadline: April 24, 2009)

**RANLP-2009**

- International Conference on Recent Advances in Natural Language Processing 2009
- Borovets, Bulgaria
- Sep 14-16, 2009 (submission deadline: abstracts: Apr 6, 2009; full submission: Apr 13, 2009; workshops: Jun 3, 2009)

**CEE-SECR 2009**

- The 5th Central and Eastern European Software Engineering Conference in Russia 2009
- Moscow, Russia
- Oct 28-29, 2009 (submission deadline: abstracts: August 7, 2009; full submission: Aug 30, 2009)

**LTC 2009**

- 4th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics
- Poznań, Poland
- November 6-8, 2009 (submission deadline: July 15, 2009)

This page was last modified 14:19, 1 July 2009. This page has been accessed 5,667 times. [Privacy policy](#) [About ACLWiki](#) [Disclaimers](#)

Powered by MediaWiki

**Figura 2.** Página de eventos de conferencias científicas del tipo “ACLWiki”

Esta página utiliza el formato “Wiki”, donde la información sobre los eventos se encuentra en forma de listas de ítems. La lista sigue un patrón común para representar la información de los CFPs. Esto se indica al comienzo de la página (ver Tabla 1 ):

URL with Short Name as a heading

- Long Name or Description
- Location
- Dates (submission date)

**Tabla 1.** Formato de datos de eventos de página tipo “ACLWiki”

Aparentemente esto facilitaría la extracción, si no fuera porque no siempre se sigue el patrón especificado.

- Existen eventos donde el nombre extendido o descripción ocupa varias líneas y es introducido como un nuevo ítem de la lista, como es el caso del evento mostrado en la Tabla 2:

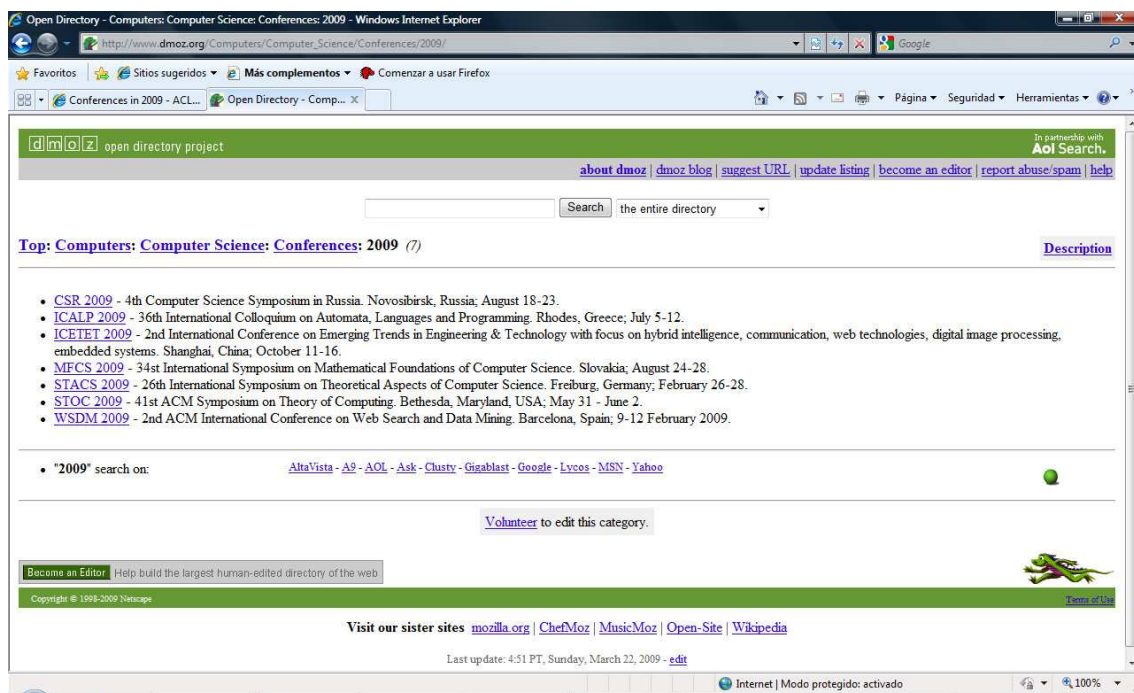
YLMP 2009

- 1st Young Linguists' Meeting in Poznan
- student conference focused on interdisciplinary linguistics
- Poznań, Poland
- April 24-26, 2009 (submission deadline: Feb 15, 2009)

**Tabla 2.** Evento mostrado en página tipo “ACLWiki”

- Requiere un parseo complicado para la obtención de fechas, puesto que estos datos están completamente desestructurados. Las fechas se pueden mostrar de múltiples formas sin seguir ningún formato común, por ejemplo:

- Jan 7-9, 2009
  - CICLing: March 1-7, 2009 (submission deadline: Nov 3 registration of abstracts, Nov 10 submission of full text). Lexicom: Feb 24-28
  - March 30-April 3, 2009 (submission deadline: Oct 10, 2008)
  - Sep 14-16, 2009 (submission deadline: abstracts: Apr 6, 2009; full submission: Apr 13, 2009; workshops: Jun 3, 2009)
- **Página de eventos “dmoz” [13]:**



**Figura 3.** Página de eventos de conferencias científicas del tipo “dmoz”

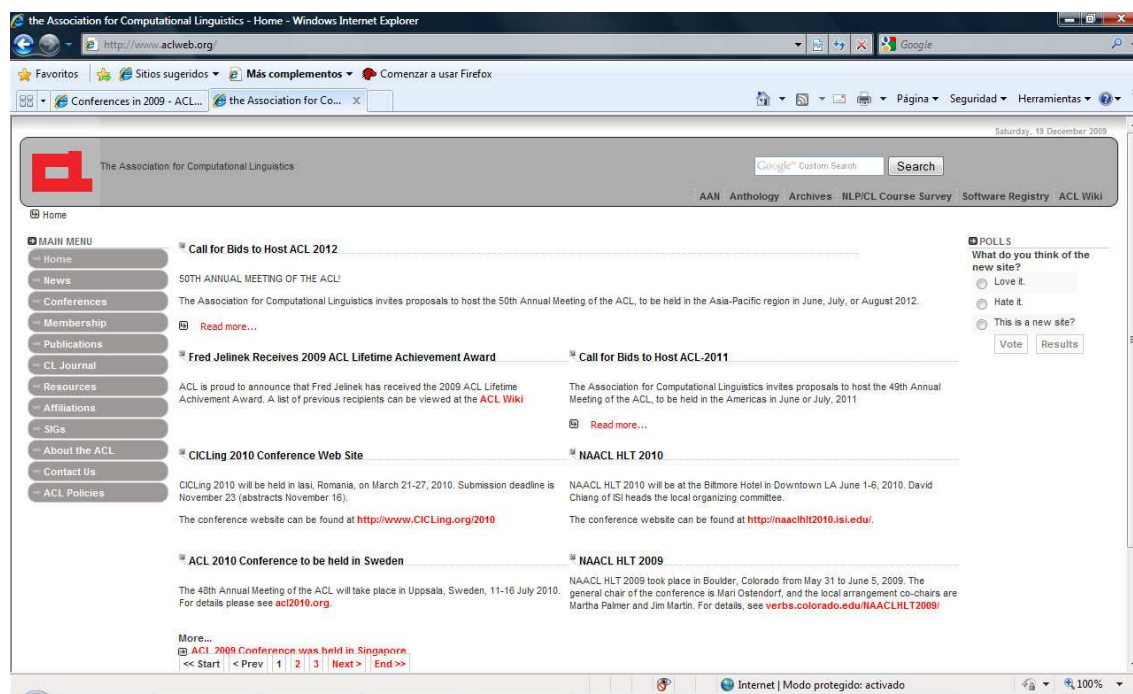
En esta página la información relacionada con los eventos también se representa en forma de lista, donde en cada ítem se encuentra la información correspondiente a cada evento en forma de texto lineal, sin estructura. El formato aparentemente que se sigue dentro de cada ítem de la lista, es el siguiente (ver Tabla 3):

Nombre corto – Descripción. Ciudad, País; Fecha

**Tabla 3.** Patrón de muestra de eventos de página tipo “dmoz”

En ningún caso se especifica el tipo de fecha que se está mostrando, ni las fechas mostradas siguen un formato común.

- **Página de eventos “aclweb” [14]:**



**Figura 4.** Página de eventos de conferencias científicas del tipo “aclweb”

En esta página la información sobre los eventos de las conferencias se muestra de forma textual, con datos totalmente desestructurados, sin seguir ningún tipo de patrón común, siendo extremadamente complicado recopilar la información de sus eventos de forma automática.

Para la realización del proceso de extracción de información de este proyecto, se han seleccionado como páginas representativas de eventos de conferencias científicas, las páginas de tipo “WikiCFP” y “ACLWiki”, puesto que son dos estilos diferentes a la

Página 20 de 237

hora de representar la información de los eventos, y además su información está bastante completa.

## **2.2 EXTRACCIÓN DE INFORMACIÓN**

La extracción de información es una disciplina dentro del procesamiento del lenguaje natural (PLN) que supone una revolución tecnológica en el ámbito de la recuperación de información y que pretende agilizar la obtención de la información útil por parte de los usuarios.

Tradicionalmente, los usuarios recuperan una gran cantidad de información y después, manualmente, deben extraer la información de estos documentos tras el análisis de los resultados recuperados. Aplicando la tecnología existente de extracción de información automática se pretende filtrar automáticamente los resultados haciendo la labor descrita anteriormente mucho menos laboriosa para los usuarios.

Los sistemas de extracción de información realizan la tarea de buscar información muy concreta en colecciones de documentos, detectar la información relevante, extraerla y presentarla en un formato susceptible de ser tratado automáticamente más tarde.

### **2.2.1 CONCEPTOS RELACIONADOS CON LA EXTRACCIÓN DE INFORMACIÓN**

#### **2.2.1.1 Diferencia entre Recuperación de Información (RI) y Extracción de Información (EI)**

La Recuperación de Información consiste en determinar cuáles son los documentos de una colección que satisfacen una “necesidad de información” de un usuario.

Mientras, la Extracción de Información se basa en localizar las porciones de texto que contengan información relevante para unas necesidades concretas de un/os usuario/s y proporcionar dicha información de forma adecuada para su proceso (de forma manual o automática).

Para poder crear un sistema de extracción de información, primero un sistema de recuperación de la información obtiene documentos con información significativa respecto a la solicitud de búsqueda y a continuación el sistema de extracción extrae y organiza la información que sea de interés.

#### **2.2.1.2 Patrones de Extracción de Información**

Dado el enorme volumen de datos en las páginas Web, es muy complicado poder recuperar la información relevante. Se consideran relevantes las páginas que son capaces de satisfacer una necesidad de información del usuario. Para que los buscadores sean capaces de recuperar páginas con información relevante han de extraer la información clave. Por lo tanto la recuperación y la extracción de la información están relacionadas, no pudiéndose recuperar páginas relevantes de las que no se hayan extraído antes los datos clave.

Los patrones pueden ser de distintos tipos:

- De léxico
- Sintácticos
- Semánticos
- De discurso

##### **a) Patrones de Léxico**

Los patrones de tipo léxico son las palabras que utilizemos para la búsqueda de información. Estas palabras se analizan por si solas y de forma independiente al contexto.

## **b) Patrones Sintácticos**

El patrón sintáctico más usado en la recuperación de la información es el llamado *part-of-speech* (POS) de una palabra, es decir, las partes del habla que son: sustantivo, verbo, artículo, adjetivo, etc.

En la recuperación de información se realiza un marcaje de "POS" cuando se asignan estas categorías gramaticales a cada palabra dada, es decir, cuando se indica la función de cada palabra en el contexto específico de la oración. Este marcaje se hace considerando características morfológicas y sintácticas del lenguaje, esto es importante ya que los nombres generalmente designan personas, lugares, cosas, y otros conceptos físicos y abstractos, los verbos suelen utilizarse para designar acciones y procesos y los adjetivos describen propiedades y estados de los nombres.

## **c) Patrones Semánticos.**

Los patrones de tipo semántico se basan en los metadatos semánticos que se añaden a la Web para describir el contenido, el significado y la relación de los datos.

En la actualidad, la *World Wide Web* está basada principalmente en documentos escritos en HTML, un lenguaje de marcas que sirve para crear hipertexto en Internet. HTML es válido para adecuar el aspecto visual del documento e incluir objetos multimedia en el texto (imágenes, esquemas de diálogo, etc.). Pero da pocas posibilidades para categorizar los elementos que configuran el texto, más allá de las típicas funciones estructurales, como sucede con otros lenguajes de maquetación. Para realizar esto, sería interesante el uso de los microformatos, como se definirá y detallará más adelante (ver apartado 2.2.3).

## **d) Patrones de discurso.**

El uso de patrones de discurso para la extracción y recuperación de información, está referido a las características de unidades de información dentro de un marco de discurso, o dentro de un marco de escritura o de estilo.



Un ejemplo muy simple sobre estos patrones de extracción es el llamado "*discourse distance*". En el proceso de extracción la distancia entre dos entidades de información es a menudo importante pues se asume que la distancia es inversamente proporcional con relación a sus significados.

### **2.2.1.3 Clasificación supervisada de sistemas de Extracción de Información**

Este tipo de extracción o recuperación de información está basado en el aprendizaje de forma supervisada. Este aprendizaje posee un algoritmo que produce una función que establece una correspondencia entre las entradas y las salidas deseadas del sistema. Un ejemplo de este tipo de algoritmo es el problema de clasificación, donde el sistema de aprendizaje trata de etiquetar (clasificar) una serie de vectores utilizando una entre varias categorías (clases). La base de conocimiento del sistema está formada por ejemplos de etiquetados anteriores.

### **2.2.1.4 Clasificación no supervisada de sistemas de Extracción de Información**

La clasificación con aprendizaje no supervisado es el aprendizaje que no necesita de un profesor, supervisor o validador externo para realizar su aprendizaje. Son capaces de modificar sus parámetros internamente, apoyándose únicamente en el conjunto de entrenamiento que requieren.

Su principal fundamento se basa en la redundancia que hay en el lenguaje natural y de esta forma poder sacar relaciones semánticas, distinguir expresiones superfluas, descubrir clases, etc.

Por otro lado, su principal ventaja frente al aprendizaje supervisado es, como se ha dicho anteriormente, que no necesitan un método de clasificación manual que haría que el sistema de recuperación o extracción de información fuera demasiado costoso. Sin embargo, se necesitan grandes cantidades de información para poder sacar las relaciones y redundancias de información anteriormente mencionadas pero este hecho no es

preocupante debido a la facilidad de recuperar, extraer y/o obtener información en la actualidad.

#### **2.2.1.5 Proyectos relacionados con la Extracción de Información**

Existen dos proyectos que proponen técnicas de extracción de información. Plantean el estudio del problema de la extracción de información desde dos ópticas distintas. Estos proyectos son:

- Proyecto ANDES [5]. Realiza la descarga de toda la información y la almacena localmente en un formato.
- Proyecto ARIADNE [6]. Mantiene la información en su lugar de origen, realizando la integración de información cada vez que se realiza una consulta al sistema.

##### **2.2.1.5.1 Proyecto ANDES**

En este proyecto se describe una metodología para extraer datos de la Web a través de técnicas de obtención de información basada en tecnologías XML. Representa un buen resumen y una serie de buenas prácticas a seguir cuando se desea afrontar un problema de estas características de forma eficiente, a través de la explicación de un software desarrollado.

En primer lugar se presentan trabajos similares, se estudian sus características y se explica que mejoras pueden realizarse sobre ello.

Posteriormente se describe paso a paso el proceso necesario:

- **Navegar por el sitio**

Se distinguen dos tipos de páginas Web: "páginas de información" con la información relevante y "páginas de navegación" que contienen enlaces. Se utiliza un robot para obtener las páginas de información y utilizar las de navegación para desplazarse, limitando el árbol de búsqueda del robot a una profundidad establecida.

- **Extracción de datos**

Antes de extraer información de las páginas es necesario formatearlas para asegurar que son bien formadas. Una vez hecho esto se le aplican varias transformaciones XSLT cuyo resultado es un documento XML que contiene sólo la información de interés. El diseño de las reglas de extracción debe ser, en la medida de lo posible, independiente de cambios en las páginas. Una aproximación a esto es evitar el uso de rutas absolutas dentro del documento, pues la probabilidad de cambio es mayor. Si se utilizan "elementos ancla" con rutas relativas la probabilidad de mantenerse estable es mayor.

- **Síntesis de hiperenlaces**

Multitud de contenidos pueden quedar ocultos en las páginas Web cuando tan sólo se utiliza la técnica de enlace a enlace para recorrerlas. Para conseguir accederlos aquí se ha realizado una serie de filtros XSLT que procesan tanto los formularios como el código JavaScript, para obtener las posibles combinaciones de enlaces que generarían peticiones correctas al gestor de contenido dinámico. Posteriormente todos esos enlaces se copian en la página original, como si de enlaces estáticos se tratara, para que el robot pueda accederlos sin tener que modificarlo.

- **Síntesis de la estructura**

Cuando se dispone de los datos y se desea introducirlos en un formato XML específico, pueden aparecer nuevas dificultades. Lo normal en un sitio Web es que encontremos la información que nos interesa en un solo bloque de texto sin etiquetar o en varios dispersos. Asimismo es posible que se proporcione alguna información de manera

gráfica (como podría ser una imagen con el precio de oferta), o que ésta se sobreentienda mirándola en la página en su conjunto. Este tipo de situaciones puede solventarse introduciendo nodos manualmente (automáticamente por XSLT), para tener la información completa en el formato final.

- **Mapeado de datos**

Una vez los datos de interés han sido separados del resto, hay que adaptarlos a una convención de nombres común, ya que probablemente varíe de un sitio a otro.

- **Integración de datos**

Por último se realiza una integración de los datos de todo el sitio en un solo documento XML.

#### **2.2.1.5.2 Proyecto ARIADNE**

Este proyecto presenta un punto de vista diferente en lo referente al almacenamiento de la información. En lugar de realizar la adquisición de las páginas y almacenar las partes interesantes en un repositorio, en este caso se considera la propia Web como base de datos. La información también se agrega, pero sólo cada vez que se realiza una consulta. Así se proporciona un método para realizar consultas similares a las de las bases de datos, realizándolas en realidad contra diversos sitios Web, dentro de un determinado dominio.

Para alcanzar el objetivo a través de esta aproximación es necesario definir:

- Un vocabulario de dominio (modelo de aplicación) en el que dar cabida tanto a la información a extraer como a las preguntas del usuario.
- Una descripción de las fuentes de información basada en ese vocabulario, cada fuente de información es definida para que sea equivalente a una clase del modelo.
- Envoltorios (*Wrappers*) para proporcionar un acceso uniforme a las fuentes de datos.
- Un planificador de consultas, que decida la forma más eficiente de servir la información.

Para explicar el problema mejor, se utiliza un ejemplo, en el que se agrega información distinta desde varios sitios. Se tienen tres sitios con información geográfica que en general proporcionan información distinta, con partes comunes. El sistema permite que se realicen consultas como "Lista de jefes de estado de todos los países cuya población es menos que diez millones", para procesarlas son divididas en partes y a cada fuente de información se le pasa la parte de la que puede encargarse.

Para construir los envoltorios del sitio Web, con el fin de proporcionar el acceso uniforme a los datos son necesarios dos modelos:

- Un modelo semántico que describa los tipos de campos disponibles en el sitio.
- Un modelo sintáctico que describa el formato de la página para que sea posible extraer cada campo.

Para obtener el modelo sintáctico se proporciona un interfaz capaz de inferir una gramática tras haber aprendido con algunos ejemplos, el número de éstos depende de la complejidad de la gramática. No se describen los algoritmos o reglas utilizadas, aunque sí se realiza una pequeña comparación contra otros sistemas similares.

Modelar la información entre distintos sitios necesita afrontar situaciones en las que tanto nombres como significados no sean estrictamente los mismos. El sistema utilizado aquí para solucionar el problema es una tabla de correspondencias, en la que se fija un sitio original, en la que un concepto toma un nombre determinado, y se indican los nombres que toma ese mismo dato en distintos sitios. Este proceso también se realiza de forma semiautomática, aunque es dependiente de un dominio concreto.

Como el enfoque seleccionado obtiene la información en tiempo real desde sitios Web que probablemente van a ser lentos, y que dependiendo del tipo de consulta puede implicar el acceso a muchas fuentes, un punto muy importante es el de la eficiencia. Para lograrla se acaba almacenando información clave localmente, tal y como se hace con las cachés.

### **2.2.2 WEB 2.0 y WEB 3.0**

En un trabajo de la Universidad ICESI del 21 de Mayo de 2009 [7], se define el término Web 2.0 como la representación de la evolución de las páginas Web tradicionales hacia las aplicaciones Web enfocadas a los usuarios. La Web 2.0 es más una actitud que una tecnología y se trata de que las aplicaciones Web generen colaboración y servicios, permitiendo que los usuarios sean autores activos de la generación de contenido.

La Web 3.0, sin embargo, trata sobre aplicaciones Web conectándose entre sí, con el fin de enriquecer la experiencia de las personas, haciendo uso de formatos especializados que amplían la información. Esto proporciona autonomía respecto del navegador y participa en la construcción de la Web Semántica. El término Web 3.0 también ha sido utilizado para describir el camino evolutivo de la red que conduce a la inteligencia artificial.

El concepto de Web Semántica llega a la actualidad a través de uno de los padres de la WWW (*World Wide Web*), el premio Nobel Sir Timothy “Tim” John Berners – Lee. La Web Semántica, trata por tanto, de poder enriquecer la comunicación mediante metadatos semánticos que aportan un valor añadido a la información, la diferencian y la hacen más inteligente.

## 2.2.3 MICROFORMATOS

### 2.2.3.1 Qué son los microformatos?

Un microformato (a veces abreviado con  $\mu$ F o uF) es una forma simple de agregar significado semántico a un contenido legible por el humano y que para la máquina es sólo texto plano. Están ideados para ser usados en páginas Web que usen HTML o XHTML, de manera tal que la información pueda ser indexada, guardada, referenciada, reutilizada o combinada.

Más técnicamente, son elementos de lenguaje de marcado, usando (X)HTML con nombres de clase específicos. En este sentido, están abiertos para que cualquiera haga uso de ellos.

Las especificaciones actuales de microformatos permiten la representación de eventos, información de contacto, relaciones sociales, direcciones, ubicaciones (coordenadas), etc.

Los microformatos son mantenidos y especificados por [Microformats.org](http://Microformats.org) [3].

Los microformatos también se podrían definir como:

- Una filosofía de publicación de contenidos en la que siempre se ha basado la Web. La información se publica para ser compartida, por lo tanto, publicar debe ser lo más fácil posible e interesa que lo publicado llegue al mayor número de lugares posibles. Esa información debe ser recogida y tratada para ser combinada por otros servicios en la Web. Los microformatos son un “atajo rápido” para la Web Semántica.
- Una técnica de construcción del código HTML basada en convenciones reconocidas por una comunidad ([microformats.org](http://microformats.org)), que permite, sin hacer desarrollos paralelos, (como en el caso del RSS, que es un archivo aparte que nosotros nunca leemos), construir minibloques con significado dentro de una



página Web, de manera que puedan ser reutilizados por otros sitios Web o aplicaciones.

### 2.2.3.2 Principios. Algunos ejemplos de utilización.

Los estándares definidos para la introducción de microformatos en (X)HTML permiten agregar significado semántico a los contenidos presentados. Esto se logra con los siguientes atributos de HTML:

- **class**
- **rel**
- **rev**

#### Ejemplo de utilización del microformato Geo:

Por ejemplo, en el texto "Las aves anidaron en 52.48,-1.89" los números se pueden entender en este contexto como coordenadas geográficas. Al rodear dicho par con etiquetas *span* (u otros elementos HTML) con nombres de clase específicos, en este caso *geo*, *latitude* y *longitude*, todos son parte de la especificación de microformato *geo*. Ver ejemplo en Tabla 4. Las máquinas pueden interpretar exactamente lo que representa cada valor y se puede realizar varias tareas, tales como: indexarlos, buscarlos en un mapa y exportarlos a un dispositivo GPS.

Las aves anidaron en
<code>&lt;span class="geo"&gt;</code>
<code>&lt;span class="latitude"&gt;52.48&lt;/span&gt;,</code>
<code>&lt;span class="longitude"&gt;-1.89&lt;/span&gt;</code>
<code>&lt;/span&gt;</code>

**Tabla 4.** Ejemplo de utilización del microformato Geo

## Ejemplo de utilización del microformato hCard:

El microformato hCard se utiliza para identificar datos personales. Por ejemplo, la información de contacto de una persona en HTML se podría representar así (Tabla 5):

```
<div>
  <div>Juan Pérez</div>
  <div>El Ejemplo S. A.</div>
  <div>604-555-1234</div>
  <a href="http://ejemplo.com/">http://ejemplo.com/</a>
</div>
```

**Tabla 5.** Ejemplo de representación de datos de contacto sin microformatos

Con el marcado del microformato hCard, esto se convierte en (Tabla 6):

```
<div class="vcard">
  <div class="fn">Juan Pérez</div>
  <div class="org">El Ejemplo S. A.</div>
  <div class="tel">604-555-1234</div>
  <a class="url" href="http://ejemplo.com/">http://ejemplo.com/</a>
</div>
```

**Tabla 6.** Ejemplo de representación de datos de contacto utilizando el microformato hCard

Aquí, el nombre formateado (*fn*), organización (*org*), número de teléfono (*tel*) y Web address (*url*) han sido identificados usando nombres de clase específicos y todo se rodea con `class="vcard"`, que indica que las otras clases forman una hCard (abreviatura para "HTML vCard"). Existen también otras clases opcionales hCard. Ahora, es posible para aplicaciones *software*, como *plugins* de navegadores, extraer la información, y transferirla a otras aplicaciones, como una agenda.

### 2.2.3.3 Ventajas del uso de microformatos

- Aventajan a la semántica basada en XML porque la información semántica puede ser intercambiada entre dos ordenadores.
- Tienen aplicación práctica inmediata.
- Son parseables inmediatamente en las páginas que los contienen.
- Son accesible a cualquier desarrollador XHTML.
- Permiten adaptar a XHTML RFCs existentes. Por ejemplo vCard pasa a ser hCard, iCal pasa a ser hCalendar, etc.
- Son compatibles con el estándar XHTML de W3C.

### 2.2.3.4 Microformatos existentes

Los siguientes microformatos han sido desarrollados para habilitar un marcado semántico para algunos tipos de información específica.

- **hAtom** (Especificación de hAtom) - para identificar información semántica en *posting* de *blogs*, también puede ser usado en artículos de noticias. (<http://microformats.org/wiki/hAtom>).
- **hCalendar** (Especificación de hCalendar) - para identificar información relacionada con eventos. (<http://microformats.org/wiki/hCalendar>).

**hCard** (Especificación de hCard) - para identificar información de contacto (<http://microformats.org/wiki/hcard>). Incluye:

- **adr** (Especificación de adr) - para direcciones postales.

- **geo** (Especificación de geo) - para coordenadas geográficas (latitud, longitud).
- **hReview** (Especificación de hReview) - para identificar críticas de artículos, servicios, etc. (<http://microformats.org/wiki/hReview>).
- **hResume** (Especificación de hResume) - para publicación de resúmenes y CVs (*Curriculum Vitae*). (<http://microformats.org/wiki/hResume>).
- **rel-directory** (Especificación de rel-directory) - para identificar enlaces en una página de un directorio. (<http://microformats.org/wiki/rel-directory>).
- **rel-nofollow** (Especificación de rel-nofollow) - advierte a los buscadores Web que no deben asignar importancia a la página a la que apunta el enlace. (<http://microformats.org/wiki/rel-nofollow>).
- **rel-tag** (Especificación de rel-tag) - indica que la página que contiene el enlace está relacionada con la etiqueta especificada. (<http://microformats.org/wiki/rel-tag>).
- **xFolk** (Especificación de xFolk) - para especificar marcadores (*bookmarks*) o enlaces etiquetados. (<http://microformats.org/wiki/xFolk>).
- **XFN** (*XHTML Friends Network*) - para especificar relaciones sociales mediante enlaces. (<http://microformats.org/wiki/xfn>).
- **XOXO** (*Extensible Open XHTML Outlines*)- para hacer esquemas con listas. (<http://microformats.org/wiki/xoxo>).
- **hRelease** - define un nuevo formato para identificar notas de prensa. (<http://www.socialtext.net/hrelease/index.cgi>).

- **unAPI** – es un pequeño API de HTTP que permite realizar operaciones básicas necesarias para realizar copias y para identificar contenidos de cualquier tipo de aplicaciones Web. (<http://unapi.info/>).
- **rel-pavatar** - Un *Personal Avatar* es un cuadro almacenado en el sitio Web de un proveedor, siguiendo unos requisitos estándares definidos en su sitio Web (<http://pavatar.com/spec/>).

El sistema *Personal Avatar* (abreviatura: Pavatar) es una forma de uso de imágenes para estandarizar este marco de perfil de proveedores, a través de varios sitios Web dejando comentario y otros contenidos. Permite a los proveedores que alojen ellos mismos sus propios *Personal Avatar*.

La mayoría de los microformatos han sido creados en el wiki de [microformats.org](http://microformats.org) y sus listas de correos, por medio del análisis de usos actuales hechos por los desarrolladores de contenidos.

Algunos otros, como rel-nofollow, unAPI, hRelease y rel-pavatar, han sido propuestos y desarrollados por otras personas.

Un detalle que se debe tener en cuenta para aprovechar un microformato es el grado de aceptación que tenga.

### 2.2.3.5 Accesibilidad de los microformatos

Los microformatos usan la etiqueta abbr (*abbreviation*) de HTML, la cual contiene un atributo *title*. En HTML en este atributo se inserta el significado de la abreviación. Dicha información es usada por algunos lectores de pantalla para "leérsela" al usuario. Sin embargo, la especificación de microformatos hace uso de la etiqueta abbr para especificar fechas, colocando en el atributo *title* la fecha en formato ISO 8601 [8], lo cual es leído como un juego de números ininteligible para las personas con discapacidad. Esto representa un fallo de accesibilidad. Este problema fue discutido por el Web Standards Project [4], en una propuesta que se denominó "hAccessibility".

### 2.2.3.6 Futuras propuestas de Microformatos

Algunos de los microformatos propuestos, que tienen mayor actividad son:

- *Citation* - para citar referencias.
- *Currency* - para expresar cantidades monetarias.
- Extensión de *Geo* - para lugares fuera de la tierra (Marte, la Luna, etc.).
- Especificación de *Species* - para nombres de seres vivos.

### 2.2.3.7 Aplicaciones que usan los microformatos

Actualmente existen extensiones para los navegadores Web que permiten tener acceso a los metadatos de los microformatos. Por ejemplo Firefox tiene una extensión llamada Operator, que permite al usuario detectar los microformatos y agregar la información que contienen a otras aplicaciones del sistema como KAddressBook o Microsoft Outlook.

En 2006, Bill Gates expresó que los microformatos son proyectos necesarios [3]

En este artículo el autor habla de la necesidad del uso de microformatos, para conseguir que todo el mundo se ponga de acuerdo, y se facilite el intercambio de datos en la Web. También considera su utilización necesaria para el uso de tarjetas de contacto, eventos, direcciones,...

## 2.2.4 EJEMPLO DE APLICACIÓN DEL MICROFORMATO HCALENDAR

Para la extracción de información relacionada con eventos, como sería necesario en este proyecto, se debería utilizar este microformato. A continuación se define más concretamente este tipo de microformato y se muestra con un pequeño ejemplo cómo se podría aplicar en una Web para facilitar la recuperación de su información.

hCalendar (abreviatura de HTML iCalendar) es un estándar de microformato para mostrar una representación semántica en (X)HTML. Se utiliza en páginas Web para representar la información relacionada con un evento de un calendario en formato iCalendar. Para ello se usan las clases de HTML y los atributos rel.

Permite que herramientas con analizadores sintácticos (como otros sitios Web, o la extensión Operator de Firefox) extraigan los detalles del evento y los muestren usando otro sitio Web, permite indexarlos, buscarlos, o cargarlos en un programa de calendarios o diarios.

Consideremos el siguiente ejemplo:

Supongamos que el texto que contiene la Tabla 7 es una parte del contenido de una página Web:

La Wikipedia en inglés fue lanzada el 15 de enero de 2001 con una fiesta desde las 2 hasta las 4pm en la casa de Jimmy Wales (más información).
---

**Tabla 7.** Texto para ejemplo de microformato

El marcado HTML para el texto anterior podría ser como se muestra en la Tabla 8:

<p> La Wikipedia en inglés fue lanzada el 15 de enero de 2001 con una fiesta desde las 2 hasta las 4pm en la casa de Jimmy Wales (<a href="http://en.wikipedia.org/wiki/History_of_Wikipedia">más información</a>) </p>
--

**Tabla 8.** Código HTML sin uso de microformato

Podemos agregar marcado hCalendar usando elementos HTML span y las clases vevent, summary, dtstart (start date), dtend (end date), locación y url como se puede ver en la siguiente tabla (Tabla 9):

```
<p class="vevent">
  La <span class="summary">Wikipedia en inglés fue lanzada</span>
  el 15 de enero de 2001 con una fiesta desde
  <abbr class="dtstart" title="2001-01-15T14:00:00+06:00">
    las 2
  </abbr>
  hasta las
  <abbr class="dtend" title="2001-01-15T16:00:00+06:00">
    4
  </abbr>
  pm en
  <span class="location">la casa de Jimmy Wales</span>
  (<a class="url"
  href="http://en.wikipedia.org/wiki/History_of_Wikipedia">más
  información</a>)
</p>
```

**Tabla 9.** Código HTML utilizando el microformato hCalendar

Nótese el uso del elemento abbr para contener formato de fecha y hora ISO-8601 para las horas de comienzo y de fin, legible para la máquina



## **2.3 APIs DE GOOGLE**

En este apartado se incluyen definiciones e información relacionada con las APIs de Google, necesarias para el tratamiento y la gestión de los eventos del sistema.

### **2.3.1 API DE DATOS DE GOOGLE (*GOOGLE DATA API*)**

Lo que viene a continuación es información genérica sobre las distintas APIs de Datos de Google, las librerías cliente que llevan asociadas e información sobre las herramientas disponibles.

Más adelante se detallará la información relacionada con el API de Datos de Google Calendar.

#### **2.3.1.1 Directorio del API**

Los siguientes servicios de Google proporcionan APIs que implementan el protocolo de datos de Google (ver Tabla 10).

Cada API tiene su propio conjunto de guías y recursos, incluyendo información sobre el uso de librerías cliente. Si se está intentando llevar a cabo una tarea con el API, la guía del desarrollador debería indicar la forma correcta. La mayoría de las APIs también incluyen código de ejemplo y otras formas sencillas para comenzar a trabajar con ellas.

API	Librerías Cliente
 Google Analytics Data Export API	JS, Java, PHP, Python, Ruby
 Google Apps APIs	Información no Disponible
 Google Base Data API	Información no Disponible
 Blogger Data API	Java, .NET, PHP, Python, JS, Obj-C
 Google Booksearch Data API	Java, PHP
 Google Calendar Data API	Java, .NET, PHP, Python, JS, Obj-C
 Google Code Search Data API	Información no Disponible
 Google Contacts Data API	Java, .NET, Python, JS, Obj-C
 Google Documents List Data API	Java, .NET, PHP, Python, Obj-C)
 Google Finance Portfolio Data API	Información no Disponible
 Google Health Data API	Java, .NET, PHP, Python, Ruby
 Google Maps Data API	Información no Disponible
 Picasa Web Albums Data API	Java, .NET, PHP, Python, Obj-C
 Google Sidewiki Data API	Java, JavaScript
 Google Sites Data API	Java, .NET, PHP, Python, Obj-C
 Google Spreadsheets Data API	Java, .NET, PHP, Python, Obj-C
 Google Webmaster Tools Data API	Información no Disponible

 YouTube Data API	Java, .NET, PHP, Python, Obj-C
--	--------------------------------

**Tabla 10.** Relación de las APIs de datos de Google con sus librerías cliente asociadas

### 2.3.1.2 Librerías cliente

Las librerías cliente para el acceso a datos de Google están disponibles para ayudar a escribir aplicaciones cliente que acceden al API.

Para cada lenguaje, la librería cliente proporciona herramientas y una capa de abstracción, permitiendo construir consultas y usar los datos de respuesta sin tener que crear peticiones HTTP o procesar respuestas HTTP a mano. Cada librería cliente proporciona clases que se corresponden con los elementos y tipos que usa.

#### 2.3.1.2.1 Librerías cliente proporcionadas por Google

Las siguientes librerías están escritas y soportadas por Google. La mayoría están alojadas en el código de Google, por eso se puede ver el código fuente (*open-source*), consultar las características, ver las preguntas más frecuentes (FAQs), y reportar errores.

- **Java client library**

Librería cliente Java proporcionada por Google.

- **.NET client library**

Librería cliente .NET proporcionada por Google y que permite el desarrollo de aplicaciones de escritorio, Web y de móviles.

- **PHP client library**

Librería cliente PHP proporcionada por Zend, como parte del *Framework* Zend. También está disponible como versión *standalone*.

- **Python pclient library**

Librería cliente Python proporcionada por Google y permite el desarrollo de aplicaciones de consola, aplicaciones Web y aplicaciones *Google App Engine*.

- **JavaScript client library**

Librería cliente JavaScript que proporciona capacidades de lectura/escritura desde dentro de JavaScript. El acceso de sólo lectura está también disponible mediante raw JSON.

- **Objective-C client library**

Librería cliente Objective-C que permite escribir aplicaciones *native Cocoa* para Mac OS X y actualmente proporciona soporte para el SDK de iPhone.

#### 2.3.1.2.2 Librerías de cliente externas

- **Force.com Toolkit** para el API de datos de Google

Esta librería cliente no está escrita ni es soportada por Google. Es la librería cliente para desarrolladores de Salesforce.com

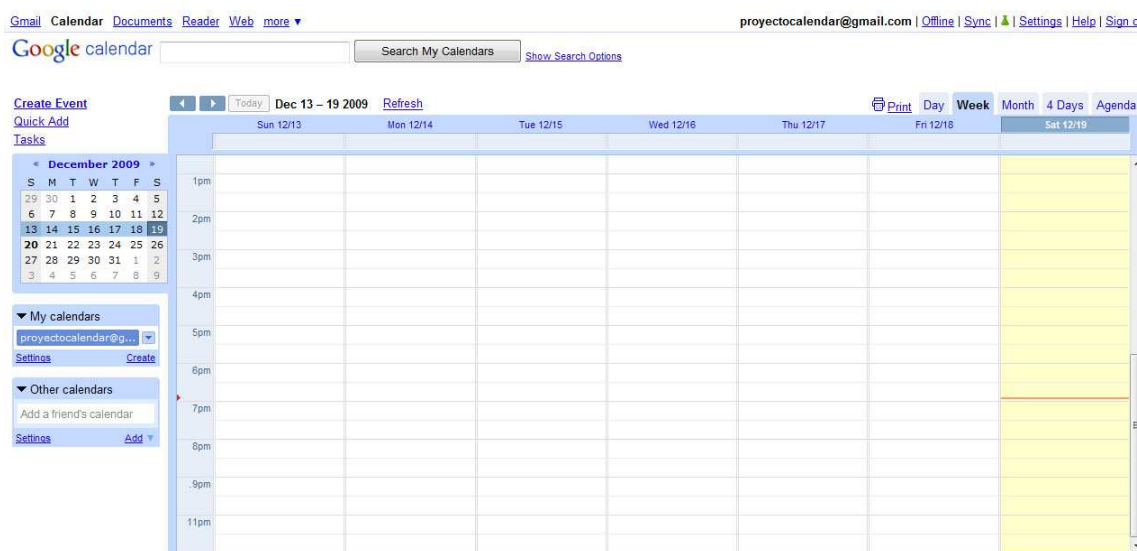
- **CalDAV API (Labs)**

CalDAV proporciona un interfaz basado en HTTP. Éste no es un API de programación, pero tiene un protocolo aparte para programas que interactúan con Google Calendar.

## 2.3.2 API DE DATOS DE GOOGLE CALENDAR

### 2.3.2.1 Descripción del API de Datos de Google Calendar

Google Calendar es una aplicación basada en Web, colaborativa, diseñada para gestionar todos los eventos importantes en un lugar. Ver Figura 5:



**Figura 5.** Imagen de Google Calendar

El API de Google Calendar y sus herramientas proporcionan a los desarrolladores un *framework* para interactuar a través de la programación con Google Calendar extendido.

Google Calendar ofrece muchas formas de crear y compartir contenido además de la interfaz Web que dispone. El protocolo de datos de Google permite a los desarrolladores crear aplicaciones que accedan a todas las funciones de Google Calendar. Se pueden insertar calendarios y eventos publicados en Google para añadir el aspecto y la funcionalidad habituales de Google a un sitio Web. También se pueden añadir marcas a la interfaz de Google y crear *gadgets* de Calendar Labs (es una parte del interfaz del sitio de Google Calendar que permite configurar algunas características del sitio) que los usuarios podrán añadir a sus propios calendarios de Google.

Google Calendar permite a las aplicaciones cliente ver y actualizar eventos del calendario como *feeds* del API de datos de Google (objetos que permiten el acceso al API de Google). La aplicación cliente puede utilizar el API de datos de Google Calendar para crear nuevos eventos, editar o eliminar eventos existentes y buscar eventos que coincidan con criterios de búsqueda determinados. Se puede utilizar esta API para: crear una interfaz de usuario Web que permita a los usuarios ver su información de Google Calendar desde tu sitio Web; añadir eventos futuros de forma programada a Google Calendar para anunciarlos; crear una aplicación para sincronizar Google Calendar y dispositivos móviles; o crear combinaciones con otras API como Google Maps.

Hay otros muchos posibles usos para el API de Datos Calendar. Por ejemplo, permite crear un *Web front end* para un calendario de grupo que usa Google Calendar como un *back end*. O permite crear un calendario público en Google Calendar para mostrar una base de datos de eventos de una compañía. También permite realizar búsquedas en calendarios relevantes para mostrar una lista de eventos en esos calendarios.

El API de Datos Calendar permite realizar la mayoría de las operaciones que un usuario podría hacer dentro del sitio de Google Calendar.

En el API de Datos además del API de Google Calendar, existen otras herramientas para trabajar con el calendario de Google, (más adelante se verán con más detalle). Estas son:

- CalDAV que proporciona acceso a la mayoría de los datos del calendario en un formato estándar de industria, apropiado para usarlo en aplicaciones de calendario tradicionales.
- El API Gadget que permite a los desarrolladores acceder directamente al interfaz Web de Google Calendar.

- Las herramientas de publicación (*Publishing Tools*) que permiten a los desarrolladores extender sus propios sitios Web con trozos de funcionalidades pre-construidos de Google Calendar.

En primer lugar, se debería decidir qué API y herramientas son las que mejor se adaptan a las necesidades y habilidades del desarrollador. Por ejemplo:

Si se necesita embeber Google Calendar en un sitio Web o permitir a los usuarios añadir un evento a sus calendarios, la solución más fácil son las herramientas de publicación (*Publishing Tools*).

El API de Datos Calendar y CalDAV proporcionan el acceso a través de la programación a Google Calendar, pero requiere experiencia en programación.

Si la aplicación se va a ejecutar dentro de Google Calendar, entonces se debería usar el API Gadget.

La tabla (Tabla 11) que se muestra a continuación, proporciona un resumen de estas opciones:

Publicación de un Evento	Acceso a Datos	Extensibilidad
Publishing Tools	Calendar Data API CalDAV	Gadget APIs

**Tabla 11.** Relación de API y herramientas con necesidades del usuario

### 2.3.2.1.1 Herramientas de publicación (Publishing Tools)

Las herramientas de publicación proporcionan elementos preconstruidos que pueden ser embebidos en un sitio Web. El uso de estas herramientas es lo más adecuado para los usuarios que están familiarizados con HTML, pero podrían no estar familiarizados con JavaScript o con programación de servidor, y que están buscando la forma más simple de embeber la funcionalidad de Google Calendar en su sitio Web. Actualmente, las herramientas que se están ofreciendo son *Embeddable Calendars* y *Event Publisher*.

- **Calendarios Embebidos (*Embeddable Calendars*)**

*Embeddable Calendars* son versiones autocontenidas de Google Calendar que pueden ser añadidas a cualquier sitio Web con algunas líneas de HTML. Estos calendarios son copias de sólo lectura del interfaz de Google Calendar, embebidas dentro de un iframe (del inglés *inline frame* o marco incorporado, es un elemento HTML que permite insertar o incrustar un documento HTML dentro de un documento HTML principal).



Permiten personalizar el color, tamaño, controles y listas de calendarios para ajustarse a las necesidades del usuario.

- **Publicador de eventos (*Event Publisher*)**

*Event Publisher* es una herramienta que crea botones que pueden ser embebidos en páginas HTML. Estos botones añaden contenido a una cuenta de Google Calendar de un usuario (eventos individuales o calendarios enteros). Los eventos de publicación no requieren tener experiencia en programación.



#### 2.3.2.1.2 Herramientas de acceso a datos

- **API de Datos Calendar (*Calendar Data API*)**

El API de Datos Calendar permite que un programa realice muchas de las operaciones disponibles a través del interfaz Web de Google Calendar. Usando este API, es posible buscar y ver los eventos públicos del calendario. Las sesiones autenticadas pueden privar el acceso a los calendarios, así como crear, editar, y eliminar tanto los eventos como los calendarios que los contienen.



Esto principalmente es para desarrolladores que tienen experiencia en



programación. Es útil para sitios o aplicaciones en los que se desea tener una integración más profunda con Google Calendar. Esta integración podría llevarse a cabo mediante una aplicación Web que crease o mostrase datos del Calendar, o podría ser una aplicación de escritorio que sincronizase un Calendar de un usuario con una aplicación de escritorio existente. Podría incluso ser un dispositivo que traspasara la experiencia del Calendar a una nueva plataforma. El API de Datos de Calendar proporciona el acceso a través de la programación a todos los datos de usuario disponibles dentro de Google Calendar. Con esto, se puede personalizar un sitio o una aplicación con la información existente de un usuario, así como realizar acciones en su nombre (como crear eventos o enviar invitaciones por mail).

Para trabajar con el API de la forma más fácil, se facilitan varias librerías cliente que abstraen el API a un modelo de objetos de un lenguaje específico. Estas librerías cliente son de código abierto y pueden ser usadas y modificadas bajo la licencia “*Apache Licence 2.0*”. Existen guías del desarrollador para Java, JavaScript, .NET, PHP, and Python, así como código de ejemplo.

- **CalDAV**

El interfaz CalDAV proporciona aplicaciones de escritorio con acceso de lectura/escritura a los datos del Calendar de los usuarios. CalDAV es un protocolo estándar de Internet para el intercambio de información de planificación de tareas entre aplicaciones cliente-servidor. Usando este protocolo, es posible almacenar y sincronizar datos de los eventos desde un calendario privado de usuario para editarlo fuera del interfaz Web de Calendar.



Esto es, principalmente, para desarrolladores que están íntimamente familiarizados con el diseño de aplicaciones de planificación, particularmente para los que usan el formato iCalendar. Otros

desarrolladores se animan a usar el API de Datos de Calendar, con el que es más fácil trabajar y proporciona funcionalidades adicionales.

### 2.3.2.1.3 Herramientas de extensibilidad (*Gadget APIs*)

Los *Gadgets* son componentes de JavaScript y HTML personalizados que se ejecutan dentro de un *iframe* y pueden ser embebidos dentro de varias páginas Web. Google Calendar incluye soporte para *Gadgets* embebidos dentro del interfaz Web de Google Calendar, creando muestras de datos de contenido enriquecido, y extendiendo Calendar con funcionalidades adicionales.

Gadgets se puede usar para crear diariamente iconos de estado, como calendarios que muestran el tiempo, contruidos dentro de Google Calendar. Estos pueden ser usados en páginas Web embebidas que se asocian con una fecha dada, como por ejemplo, un calendario que muestra escenas de una película, donde se muestran los *trailers* dentro de la descripción de cada evento. También pueden ser usados para añadir nuevos controles, como las características “*Jump To Date*” (Ir a Fecha) y “*Next Meeting*” (Próxima Cita) de Calendar Labs.

Google Calendar proporciona dos APIs para Calendars embebidos:

- **Event Gadgets**

*Event Gadgets* permite embeber contenidos Web dentro de un calendario. Este calendario puede ser específico para un único usuario, o compartido por varios usuarios. Estos eventos pueden aparecer como eventos normales, eventos con iconos, o iconos independientes sin evento asociado. El *gadget* cargará una URL específica dentro de un *iframe* en lugar de encapsularlo como un evento normal.



*Event gadgets* es útil cuando se importan eventos dentro de Google Calendar, y esos eventos tienen información adicional almacenada fuera del

Calendar, en la que el usuario puede estar interesado. Algunos ejemplos serían datos del tiempo, recordatorios de cumpleaños, estrenos de películas, y otros. Sin embargo, ofrecen soporte limitado de JavaScript, y no proporcionan acceso para controlar el interfaz de usuario de Calendar.

- **Sidebar Gadgets**

Sidebar Gadgets son pequeñas aplicaciones Web portables que se encuentran en la parte izquierda del interfaz de usuario de Calendar. Soportan Gadget APIs, y son una excelente elección cuando se implementan funcionalidades complejas. También tienen acceso a datos de eventos y controles del interfaz de usuario de Calendar, lo cual también les permite actuar como *plugins* para el interfaz de Google Calendar.



Muchas de las características de Calendar Labs, son implementadas como *Sidebar Gadgets*, incluyendo “*Jump To Date*” (Ir a Fecha), “*Next Meeting*” (Próxima Cita) y “*World Clock*” (Reloj Mundial). Adicionalmente, existen *Gadgets* de otros sitios que pueden ser convertidos dentro de *Sidebar Gadgets* con pocas o ninguna modificación.

### 2.3.2.2 Autenticación del servicio Calendar

La inserción, actualización o eliminación de entradas, requieren una autenticación usando uno de los dos sistemas de autenticación diferentes soportados. El método apropiado de autenticación depende del tipo de cliente que se vaya a utilizar.

Si el cliente es un cliente instalado (*standalone*) de un sólo usuario (como una aplicación de escritorio), entonces se debe usar el sistema ***ClientLogin***. Si el cliente es una aplicación Web multi-usuario, se debería usar un sistema ***Authsub***. Ambos métodos implican interactuar con un servicio de autenticación. El servicio de autenticación

devuelve un token de autenticación que el cliente puede enviar al servicio del API de Datos Calendar con cada petición de subsecuencia procedente del usuario.

### 2.3.2.2.1 Autenticación AuthSub proxy

La autenticación del proxy *AuthSub* se usa en aplicaciones Web que necesitan autenticar a sus usuarios en cuentas de Google. El operador del sitio Web no necesita acceder al usuario y contraseña del usuario, sólo se requieren tokens AuthSub especiales.

Para adquirir un token AuthSub para un usuario dado, la aplicación debe redirigir al usuario a la URL AuthSubRequest, mediante la cual puede logarse en su cuenta de Google.

La URL AuthSubRequest podría parecerse a la siguiente (Tabla 12):

<https://www.google.com/accounts/AuthSubRequest?scope=http%3A%2F%2Fwww.google.com%2Fcalendar%2Ffeeds%2F&session=1&secure=0&next=http%3A%2F%2Fwww.coolcalendarsite.com%2Fwelcome.html>

**Tabla 12.** Url de autenticación AuthSubRequest

La descripción de los parámetros enviados en la URL AuthSubRequest, es la siguiente (Tabla 13):

Parámetro	Descripción
<b>Scope</b>	URL base para peticiones <i>feed</i> . Para Google Calendar <i>feeds</i> este valor es: <a href="http://www.google.com/calendar/feeds/">http://www.google.com/calendar/feeds/</a>
<b>Session</b>	Indica si el token devuelto puede ser intercambiado por un token multiuso (session).
<b>Secure</b>	Indica si el token devuelto será un token seguro.
<b>Next</b>	La URL de la página a la que Google debería redirigir al usuario después de la autenticación.

**Tabla 13.** Parámetros de la URL de autenticación AuthSubRequest

Después de que los usuario se han autenticado, el sistema AuthSub redirige a la URL especificada en el parámetro *next* de la consulta de la URL AuthsubRequest. El sistema añade un *token* de autenticación a esa URL (Tabla 14). Este valor de token representa un *token* AuthSub de un sólo usuario.

<code>http://www.coolcalendarsite.com/welcome.html?token=yourAuthToken</code>
---

**Tabla 14.** Url de autenticación AuthSubRequest después de que el usuario se haya autenticado

#### 2.3.2.2.2 Autenticación ClientLogin username/password

El método ClientLogin requiere tener acceso a un usuario y una contraseña de una cuenta Google. Para recibir el *token* de autenticación usando el mecanismo ClientLogin hay que enviar una petición POST a la siguiente URL (Tabla 15):

<code>https://www.google.com/accounts/ClientLogin</code>
--

**Tabla 15.** URL de petición POST para autenticación ClientLogin

El cuerpo de POST debería contener un conjunto de parámetros, como se describe en la Tabla 16. Deberían parecerse a los parámetros pasados a un formulario HTML, usando el tipo de contenido: `application/x-www-form-urlencoded`.

Parámetro	Descripción
<b>Email</b>	El mail del usuario.
<b>Passwd</b>	La contraseña del usuario.
<b>Source</b>	Identifica la aplicación cliente. Debería tomar la forma: companyName-applicationName-versionID Debajo, se usará el nombre: exampleCo-exampleApp-1.
<b>service</b>	La cadena cl, es el nombre del servicio para Google Calendar.

**Tabla 16.** Parámetros del cuerpo de la petición POST para autenticación ClientLogin

Si la petición de autenticación falla, se recibirá un código de estado: “HTTP 403 Forbidden”. Si es exitosa, la respuesta del servicio es un código de estado: “HTTP 200 OK”, más tres códigos alfanuméricos en el cuerpo de la respuesta: SID, LSID, y Auth. El valor Auth es el *token* de autorización que enviará a Calendar con su petición, por eso mantiene una copia de ese valor. Los valores de SID y LSID se pueden ignorar.

Desde que todas las peticiones a *feeds* privados (accesos a entradas de Calendars privados) requieren autenticación, hay que establecer una cabecera Authorization en la petición, con el siguiente formato (Tabla 17):

Authorization: GoogleLogin auth=yourAuthToken
---

**Tabla 17.** Cabecera de autorización

Donde yourAuthToken es la cadena Auth devuelta por la petición de autenticación.

Este sistema de autenticación (conocido como “Autenticación Google para aplicaciones instaladas; también conocido como ClientLogin), es apropiado sólo para el uso en aplicaciones clientes instaladas tales como clientes de escritorio, no para aplicaciones Web. Para Aplicaciones Web, se debería usar en su lugar el *proxy* de autenticación de cuenta o AuthSub.

### 2.3.2.2.3 Autenticación Magic Cookie

La URL de *feed* (acceso a la entrada de un Calendar) más simple es la que se usa en calendarios de sólo-lectura, el *feed* privado “*Magic Cookie*”, cuya URL no requiere autenticación. El procedimiento usual para determinar la URL, implica usar un navegador con javascript habilitado para conseguir la URL manualmente. Si no se puede o no se quiere usar un navegador, entonces, en su lugar, se puede interactuar con Calendar usando uno de las siguientes *feeds* de URLs, que requieren autenticación, pero que se pueden construir sin necesidad de un navegador.

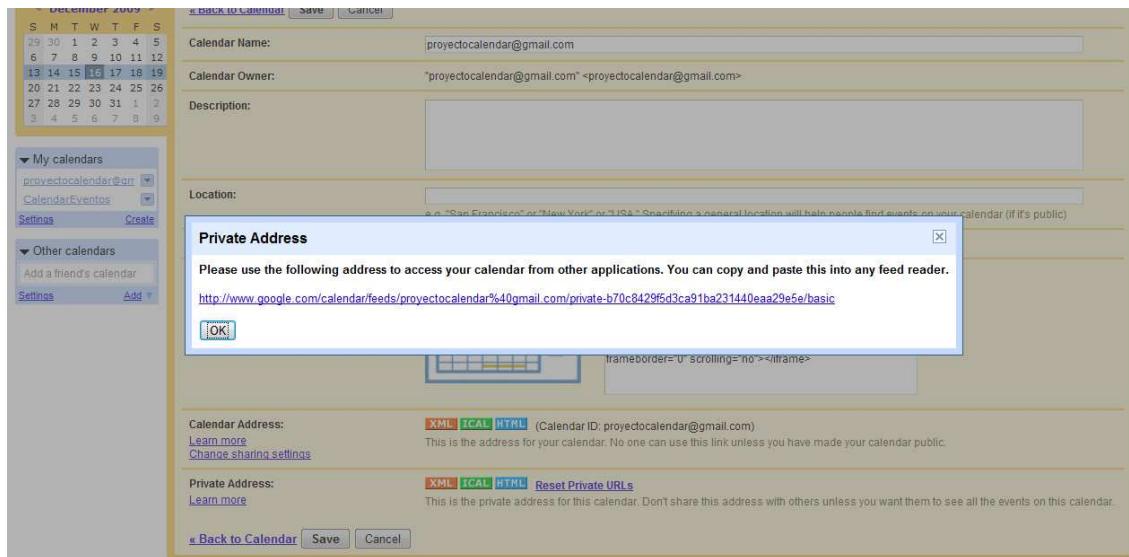
Para encontrar la URL del *feed* de “*Magic Cookie*” de Calendar:

- En la lista de calendarios en lado izquierdo de la página principal de Google Calendar, encontrar el calendario con el que se quiere interactuar. Se puede crear un nuevo calendario para este propósito si se quiere.
- Hacer clic con el botón de fecha a la derecha del calendario. Seleccionar “Calendar Settings” del menú desplegable. Aparecerá la página de “Calendar Details” (ver Figura 6).

The screenshot shows the Google Calendar interface. On the left, there's a sidebar with a calendar view for December 2009 and a list of calendars including 'proyectocalendar@gmail.com'. The main area displays the 'Calendar Details' for 'proyectocalendar@gmail.com'. It includes fields for 'Calendar Name', 'Calendar Owner', and 'Description'. There's a 'Location' field with a hint about specifying a general location. The 'Calendar Time Zone' is set to 'GMT+01:00 Madrid'. An 'Embed This Calendar' section provides an embed code for websites. At the bottom, there are 'Calendar Address' and 'Private Address' sections with links to XML, iCal, and HTML formats. Navigation buttons like 'Back to Calendar', 'Save', and 'Cancel' are present.

**Figura 6.** Pantalla de Google Calendar Details para el Calendar “proyectoCalendar@gmail.com”

- Bajar el *scroll* a la sección “Private Address” y seleccionar el botón XML. Aparecerá la URL del *feed*. (ver Figura 7).



**Figura 7.** URL “Magic Cookie” para el Calendar seleccionado

- Copiar la URL. Ésta es la URL del *feed* privado “Magic Cookie” de sólo-lectura del calendario. Incluye una cadena codificada especial que permite leer el *feed* privado sin necesitar autenticación. Será la URL que se usará para la petición de un *feed* desde Calendar, por eso no habrá que autenticarse para ver el *feed*.

La URL del *feed* de la Figura 7 tiene la siguiente forma (ver Tabla 18):

<code>http://www.google.com/calendar/feeds/userID/private-magicCookie/basic</code>
--

**Tabla 18.** URL “Magic Cookie”

Donde *magicCookie* es un código especial que permite leer el *feed* privado sin tener que autenticarse.

Bajo algunas circunstancias, se podría querer generar una nueva URL “Magic Cookie”. (Hay que tener en cuenta que si alguien más consigue acceder a esa URL, podrían ver el



calendario sin autenticación). Para realizar esto, hay que acceder desde el enlace “Reset Private URL” en la página “Calendar Details”.

### 2.3.3 API AJAX DE GOOGLE

El API AJAX (*Asynchronous JavaScript And XML*) de Google contiene librerías que permiten crear sitios Web completos y dinámicos utilizando JavaScript y HTML. Por ejemplo, se puede insertar un mapa o un cuadro de búsqueda dinámico en un sitio Web, o descargar *feeds* con código JavaScript.

A continuación se muestran los tipos de APIs que hay junto con una breve descripción de lo que permite cada uno:

- **API de Google Maps**

Permite añadir mapas a una página Web (Figura 8).



**Figura 8.** Representación de mapa de Google

- **API AJAX de Google para búsquedas**

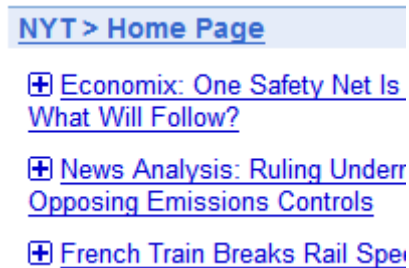
Permite integrar búsquedas de Google en un sitio Web, incluido Google Videos, Google Maps y Google Noticias (Figura 9).



**Figura 9.** Representación de buscador de Google

- **API AJAX de Google para feeds**

Permite descargar *feeds* RSS y Atom de cualquier servidor usando JavaScript (Figura 10).



**Figura 10.** Representación de *feeds* de Google

- **API de visualización de Google**

Permite crear aplicaciones de elaboración de informes y visualizaciones con acceso a datos estructurados en un formato común (Figura 11).



**Figura 11.** Representación de informes de Google

- **API AJAX de idiomas de Google**

Permite traducir y detectar fácilmente una gran variedad de idiomas utilizando JavaScript (Figura 12).



**Figura 12.** Representación de idiomas de Google

- **API de bibliotecas AJAX**

Permite acceder a bibliotecas AJAX populares directamente desde los servidores de Google (Figura 13).



**Figura 13.** Representación de bibliotecas AJAX

- **API de Google Libros**

Permite integrar de una forma sencilla en un sitio Web el contenido y la vista previa de libros (Figura 14).



**Figura 14.** Representación del API de Google para libros

- **API de Google Earth**

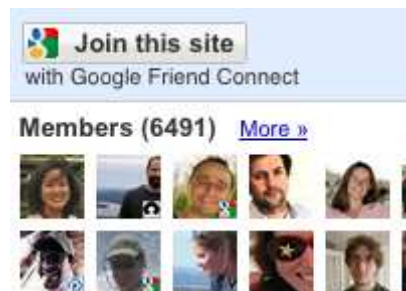
Permite añadir funcionalidades de Google Earth a una página Web (Figura 15).



*Figura 15.* Representación de Google Earth

- **API de Google Friend Connect**

Permite integrar flujos y datos sociales directamente en el marcado de una página Web (Figura 16).



*Figura 16.* Representación de Google Friend Connect

## 2.3.4 API DE GOOGLE MAPS

El API de Google Maps permite insertar Google Maps en las páginas Web utilizando JavaScript. El API proporciona varias utilidades para manipular mapas y añadir contenido a los mismos mediante algunos servicios, permitiendo crear sólidas aplicaciones de mapas en un sitio Web. Para utilizar el API, primero hay que registrarte para obtener una clave de API.

En los siguientes subapartados se expone brevemente cómo se puede integrar Google Maps en una aplicación.

Para seguir la documentación que se facilita en Google sobre Google Maps, el desarrollador deber estar familiarizado con programación JavaScript y conceptos de programación relacionados con objetos. También debe estar familiarizado con Google Maps desde el punto de vista del usuario.

#### **2.3.4.1 Creación de mapplets de Google**

El API de Google Maps admite mapplets (mapas + gadgets), que permiten insertar en Google Maps aplicaciones alojadas en ubicaciones externas. Estos mapplets se ejecutan en sus propios contenedores (*iFrames*), lo que permite, combinar código de un sitio con el de otro. Esto abre un panorama completamente nuevo para la programación en Google Maps. Crear mapplets es sencillo, aunque hay algunas cosas que no se hacen del mismo modo que con el API estándar de Google Maps.

#### **2.3.4.2 Uso del cargador de API AJAX de Google**

El API de Google Maps ahora está completamente integrado en las APIs AJAX de Google. Esta estructura permite cargar una clave de API para todas las API AJAX de Google admitidas (incluida la de Google Maps) y ofrece también un espacio de nombre común para cada API, lo que permite un funcionamiento conjunto de distintas APIs de Google. Si se decide no utilizar el API AJAX de Google, se puede seguir utilizando el espacio de nombre existente.

El uso del API AJAX de Google es relativamente sencillo. Si se quiere cambiar una aplicación para utilizar esta estructura, se deben seguir los siguientes pasos:

- En lugar de cargar el API desde: *http://maps.google.com/apis*, hay que cargar el cargador común desde: *http://www.google.com/jsapi*. Se puede incluir la clave de API de Google Maps en la siguiente URL (Tabla 19):

```
<script type="text/javascript" src="http://www.google.com/jsapi?key=ABCDEFGH"></script>
```

**Tabla 19.** Url que contiene la clave del API de Google

- Se debe cargar el API específica que se quiera utilizar con el método *google.load*. Este método utiliza un argumento con el API y el número de versión específico que se va a cargar (Tabla 20):

```
<script type="text/javascript">
  google.load("maps", "2");
</script>
```

**Tabla 20.** Ejemplo de utilización del método *google.load*

- Hay que utilizar el espacio de nombre *google.maps.\** para todos los métodos, las clases y las propiedades que se utilicen actualmente en el API de Google Maps, sustituyendo el prefijo *G* por este espacio de nombre. La inicialización de los objetos se realizan con *google.setOnLoadCallback()*.

#### 2.3.4.3 Selección del idioma del API de Google Maps

El API de Google Maps utiliza la configuración de idioma preferido del navegador para mostrar información textual como, por ejemplo, nombres de controles, avisos de copyright e instrucciones de circulación. Si se quiere cambiar la configuración de idioma del navegador en el que el API de Google Maps muestra la información y obligarlo a utilizar un idioma concreto, se puede añadir un parámetro *hl* opcional a la etiqueta *<script>* al incluir el código JavaScript del API de Google Maps especificando el idioma de dominio que se desea utilizar.

Por ejemplo, si se quiere que la información del API de Google Maps aparezca en alemán, habría que añadir *&hl=de* a la etiqueta *<script>* del siguiente modo (Tabla 21):

```
<script type="text/javascript"
src="http://maps.google.com/maps?file=api&v=2&key=abcdefg&hl=de"
```

**Tabla 21.** Ejemplo de selección de idioma del API de Google Maps

#### 2.3.4.4 Especificación del parámetro sensor

Si se utiliza el API de Google Maps, ahora se debe indicar si la aplicación utiliza un dispositivo de detección (sensor) (por ejemplo, un localizador GPS) para determinar la ubicación del usuario. Esto resulta de especial importancia para los dispositivos móviles. Las aplicaciones deben transmitir un parámetro *sensor* obligatorio a la etiqueta *<script>* cuando se incluya el código JavaScript del API de Google Maps, indicando si la aplicación está utilizando un dispositivo de detección.

Las aplicaciones que determinan la ubicación del usuario a través de un sensor deben transmitir *&sensor=true* cuando se cargue el código JavaScript del API de Google Maps a través de <http://maps.google.com/maps> o con el cargador común.

Hay que tener en cuenta que aunque el dispositivo no utilice un dispositivo de detección, se debe seguir transmitiendo este parámetro, estableciendo su valor en *false*.

#### 2.3.4.5 Compatibilidad

Antes de empezar a explorar el API de Google Maps, se deben tener en cuenta las siguientes cuestiones para asegurar que la aplicación funcione sin problemas en las plataformas en las que se vaya a utilizar:

- **Compatibilidad del navegador:**



El API de Google Maps admite los mismos navegadores que el sitio Web de Google Maps.

- **La secuencia de comandos:**

*http://maps.google.com/maps?file=api&v=2*, se puede analizar sin errores en casi todos los navegadores, así que se puede incluir con seguridad esa secuencia de comandos antes de comprobar la compatibilidad.

En algunas ocasiones, las diferentes aplicaciones requieren distintos comportamientos cuando los usuarios utilizan navegadores incompatibles. El API de Google Maps ofrece un método global (*GBrowserIsCompatible()*) de comprobación de la compatibilidad, pero no dispone de ningún comportamiento automático para los casos en los que detecta un navegador incompatible.

#### **2.3.4.6 XHTML y VML**

Se recomienda el uso de XHTML conforme con los estándares en las páginas que contienen mapas. Cuando los navegadores ven el código XHTML DOCTYPE en la parte superior de la página, muestran la página en "modo conforme con los estándares", lo que hace que el diseño y los comportamientos en los navegadores sean mucho más predecibles. Si las páginas no incluyen esa definición se pueden ocasionar inconsistencias en el diseño.

Si se quieren mostrar polilíneas en un mapa (como las líneas que utiliza Google Maps para las instrucciones para llegar en coche), hay que incluir el espacio de nombre VML en el documento XHTML para que todo funcione correctamente en Internet Explorer. El principio del documento XHTML debe ser como el que se indica en la Tabla 22:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xmlns:v="urn:schemas-microsoft-
com:vml">
```

**Tabla 22.** Ejemplo de comienzo de un documento XHTML

### 2.3.4.7 Actualizaciones del API

En la siguiente URL: *http://maps.google.com/maps?file=api&v=2*, el parámetro “v” hace referencia al número de *versión* del API de Google Maps que se va a utilizar. La mayoría de los usuarios del API probablemente prefieran utilizar la “versión 2” actual del API especificando el parámetro *v=2* en esa URL. Si se prefiere, se puede utilizar *v=2.x* para utilizar la última versión (con las funciones más recientes). No obstante, se debe tener en cuenta que la última versión puede no ser tan fiable como la versión 2.

También se puede utilizar una versión “estable” del API con el parámetro *v=2.s*. Esta versión se actualiza con menos frecuencia (cada pocos meses). Dado que las versiones 2 y 2.x se actualizan aproximadamente cada dos semanas, algunos desarrolladores pueden preferir utilizar la versión estable 2.s. Se ha de tener en cuenta que la versión estable puede ser varias versiones anterior a la actual y no incluir, por tanto, algunas de las últimas funciones.

Cada versión del API se etiqueta al actualizarla (por ejemplo, “versión 2.76”). Cuando se actualiza el API, en algunas ocasiones el código anterior basado en versiones previas puede no funcionar según lo previsto. Si se necesita vincular inevitablemente la aplicación a una determinada versión del API, se puede hacer incluyendo esa versión de forma explícita en el parámetro *v* (por ejemplo, *v=2.75*). Sin embargo, esta opción no es recomendable ya que siempre será mejor que basar los productos en el código que se actualice.

Si se realiza una actualización significativa del API en el futuro, se cambiará el número de versión principal y se publicará un aviso en Google Code y en el foro de

Página 66 de 237

desarrolladores del API de Google Maps. Cuando esto suceda, la idea es mantener ambas versiones disponibles durante un tiempo para que se pueda migrar el código.

El equipo de Google Maps también actualiza de forma invisible el API con las mejoras de rendimiento y las correcciones de errores más recientes. Estas correcciones sólo deberían mejorar el rendimiento y solucionar errores, pero es posible que rompan de forma inadvertida algunos clientes del API. Se informa de estos problemas en el foro de desarrolladores del API de Google Maps.

### 2.3.5 API AJAX DE GOOGLE SEARCH

El API AJAX de Google para búsquedas es una biblioteca JavaScript que permite insertar Google Search en páginas Web y otras aplicaciones Web. Para entornos Flash y otros entornos que no utilizan JavaScript, el API muestra una interfaz que devuelve resultados con codificación JSON (*JavaScript Object Notation*) que son fácilmente procesados por la mayoría de los lenguajes y tiempos de ejecución.

El API AJAX de Google para búsquedas proporciona objetos Web simples que realizan búsquedas integradas en varios servicios de Google (búsqueda en la Web, local, de vídeos, de blogs, de noticias, de libros (experimental), de imágenes y de patentes). Si la página Web está diseñada para ayudar a los usuarios a crear contenido (por ejemplo, tableros de mensajes, *blogs*, etc.), el API está diseñado para realizar estas actividades, ya que permitirá copiar los resultados de búsqueda directamente en sus mensajes.

Para poder insertar Google Search en una página Web personal, hay que registrarse para obtener una clave de API AJAX para búsquedas. Esta clave es válida para un único directorio del servidor Web (y los subdirectorios correspondientes). Si se registra, por ejemplo, la URL: *http://www.mygooglesearchssite.com/mysite*, se creará una clave que se podrá utilizar con todas las URL del directorio *http://www.mygooglesearchssite.com/mysite/*

Para obtener una clave de API de Google, el usuario debe tener una cuenta de Google. La clave de API se vinculará directamente a la cuenta de Google. Para varios sitios Web, se pueden obtener varias claves de API para la misma cuenta.

#### 2.3.5.1 Compatibilidad del navegador

El API AJAX de Google para búsquedas es compatible actualmente con Firefox 1.5 o superior, Internet Explorer 6, Safari, Opera 9 o superior y Chrome.

A través del método *google.load* se carga el API AJAX de Google para búsquedas. El método *google.load* utiliza un argumento con el API y el número de versión específico que se va a cargar.

### **2.3.5.2 Actualizaciones del API**

Como se comentaba anteriormente, el segundo argumento de la instrucción: *google.load('search', '1.0');* contiene el nivel de versión esperado de este API (en este caso, la versión "1.0"). Cuando se realice una actualización significativa del API, se aumentará la versión un número y se publicará un aviso en el grupo de debate del API AJAX de Google para búsquedas. Se recomienda anotar cualquier cambio de código necesario cuando esto ocurra y actualizar las direcciones URL a la nueva versión cuando sea oportuno.

Después de lanzar una nueva versión, Google admite la versión antigua y la nueva de forma simultánea durante un período de tiempo limitado, que se prevé será de varios meses. Una vez que este período finalice, ya no se aceptarán las solicitudes de los clientes que utilicen el API antigua, por lo que se debe cambiar el código inmediatamente tras conocer que existe una nueva versión.

El equipo del API AJAX de Google para búsquedas también actualiza periódicamente el API con las mejoras de rendimiento y las correcciones de errores más recientes que no requieren una actualización de versión. En la mayoría de los casos, estas correcciones deben ser transparentes para el usuario, pero es posible que rompan por error algunos clientes del API. También en este caso se informa de estos problemas a través del grupo de debate del API AJAX de Google para búsquedas.

### 2.3.5.3 Aspectos generales del API

El API AJAX de Google para búsquedas está formado por varias clases de objetos:

- ***google.search.SearchControl***: esta clase proporciona al usuario la interfaz y la coordinación sobre una serie de objetos buscadores, donde cada objeto buscador está diseñado para realizar búsquedas y devolver una clase determinada de resultados (búsqueda en la Web, búsqueda local, etc.).
- ***google.search.Search***: esta clase base es la clase desde la que heredan todos los "buscadores". Define la interfaz que deben implementar todos los servicios de búsqueda.
- ***GResult***: esta clase base incluye los resultados de búsqueda producidos por los objetos buscadores.
- ***google.search.SearcherOptions***: esta clase configura el comportamiento de los objetos buscadores cuando se añade a un control de búsqueda.

## **2.4 HERRAMIENTAS Y TECNOLOGÍAS QUE SE UTILIZAN EN EL SISTEMA**

A continuación se muestra información referente a las herramientas y tecnologías relacionadas con el desarrollo de este proyecto.

### **2.4.1 MICROSOFT VISUAL STUDIO**

Microsoft Visual Studio es un entorno de desarrollo integrado (IDE, por sus siglas en inglés) para sistemas operativos Windows. Soporta varios lenguajes de programación tales como Visual C++, Visual C#, Visual J#, ASP.NET y Visual Basic .NET, aunque actualmente se han desarrollado las extensiones necesarias para muchos otros.

Visual Studio permite a los desarrolladores crear aplicaciones, sitios y aplicaciones Web, así como servicios Web en cualquier entorno que soporte la plataforma .NET (a partir de la versión net 2002). Así se pueden crear aplicaciones que se intercomunicuen entre estaciones de trabajo, páginas Web y dispositivos móviles.

#### **Visual Studio 2005**

Visual Studio 2005 se empezó a comercializar a través de Internet en Octubre de 2005 y llegó a los comercios a finales de mes, en inglés. Su versión en castellano no se comercializó hasta el 4 de Febrero de 2006.

Características:

- Incluye tipos genéricos, similares en muchos aspectos a las plantillas de C#.
- Incluye un diseñador de implantación, que permite que el diseño de la aplicación sea validado antes de su implantación.

- Incluye un entorno para publicación Web y pruebas de carga para comprobar el rendimiento de los programas bajo varias condiciones de carga.
- Añade soporte de 64-bit. Aunque el entorno de desarrollo sigue siendo una aplicación de 32 bits Visual C++ 2005 soporta compilación para x86-64 (AMD64 e Intel 64) e IA-64 (Itanium). El SDK incluye compiladores de 64 bits así como versiones de 64 bits de las librerías.

Visual Studio 2005 tiene varias ediciones radicalmente diferentes entre sí: Express, Standard, Professional, Tools for Office, y 5 ediciones Visual Studio Team System. Éstas últimas se proporcionaban conjuntamente con suscripciones a MSDN cubriendo los 4 principales roles de la programación: profesionales relacionados con la arquitectura, desarrolladores de Software, de pruebas, y de Bases de Datos. La funcionalidad combinada de las 4 ediciones Team System se ofrecía como la edición Team Suite.

## 2.4.2 EXPRESIONES REGULARES

Una expresión regular, a menudo llamada también patrón, es una expresión que describe un conjunto de cadenas sin enumerar sus elementos.

Por ejemplo, el grupo formado por las cadenas *Handel*, *Händel* y *Haendel* se describe mediante el patrón "H(a|ä|ae)ndel".

Una expresión regular es una forma de representar a los lenguajes regulares (finitos o infinitos) y se construye utilizando caracteres del alfabeto sobre el cual se define el lenguaje.



## Operaciones básicas:

### 1. Alternación

Mediante una barra vertical se separan distintas alternativas. Por ejemplo, la expresión regular "marrón|castaño" recogería las cadenas *marrón* o *castaño*.

### 2. Cuantificación

Un cuantificador tras un carácter especifica la frecuencia con la que éste puede ocurrir. Los cuantificadores más comunes son +, ? y \*:

- Cuantificador +: Indica que el carácter al que sigue debe aparecer al menos una vez. Por ejemplo, "ho+la" describe el conjunto infinito *hola*, *hoola*, *hoolola*, *hoooola*, etcétera.
- Cuantificador ?: Indica que el carácter al que sigue puede aparecer como mucho una vez. Por ejemplo, "ob?scuro" casa con *oscuro* y *obscurro*.
- Cuantificador \*: Indica que el carácter al que sigue puede aparecer cero, una, o más veces. Por ejemplo, "0\*42" casa con *42*, *042*, *0042*, *00042*, etcétera.

### 3. Agrupación

Los paréntesis pueden usarse para definir el ámbito y precedencia de los demás operadores. Por ejemplo, "(p|m)adre" es lo mismo que "padre|madre", y "(des)?amor" coincidiría con *amor* y *desamor*.

Los constructores pueden combinarse libremente dentro de la misma expresión, por lo que "H(ae?|ä)ndel" equivale a "H(a|ae|ä)ndel".

La sintaxis precisa de las expresiones regulares cambia según las herramientas y aplicaciones consideradas.

Su utilidad más obvia es la de describir un conjunto de cadenas, lo que resulta de utilidad en editores de texto y aplicaciones para buscar y manipular textos. Muchos lenguajes de programación admiten el uso de expresiones regulares con este fin. Por ejemplo, Perl tiene un potente motor de expresiones regulares directamente incluido en su sintaxis. Las herramientas proporcionadas por las distribuciones de Unix fueron las primeras en popularizar el concepto de expresión regular.

### 2.4.3 REGULATOR 2.03

Regulator es una herramienta avanzada y gratuita para la validación y aprendizaje de expresiones regulares. Permite construir y verificar una expresión regular contra cualquier entrada de texto, fichero o página Web, y mostrar las coincidencias o reemplazos resultantes en un árbol jerárquico fácil de entender.

Las acciones de expresiones regulares soportadas son las siguientes:

- Encontrar múltiples coincidencias.
- Reemplazar texto de acuerdo a una expresión.
- Separar un texto de acuerdo a una expresión.

Regulator también está integrada con [RegexLib.com](http://RegexLib.com), una biblioteca gratuita de expresiones regulares, ayudando a encontrar expresiones regulares listas para ser utilizadas directamente.

### 2.4.4 HTML

El lenguaje de marcas de hipertexto, HTML o (*HyperText Markup Language*) se basa en el metalenguaje SGML (*Standard Generalized Markup Language*) y es el formato de los documentos de la *World Wide Web*. El *World Wide Web Consortium* (W3C) es la organización que desarrolla los estándares para normalizar el desarrollo y la expansión de la Web y la que publica las especificaciones relativas al lenguaje HTML.

HTML fue concebido como un lenguaje para el intercambio de documentos científicos y técnicos adaptado para su uso por no especialistas en tratamiento de documentos. HTML resolvió el problema de la complejidad de SGML sirviéndose de un reducido conjunto de etiquetas estructurales y semánticas apropiadas para la realización de documentos relativamente simples. Pero, además de simplificar la estructura de los documentos, HTML soportaba el hipertexto.

El lenguaje HTML nace en 1991 de manos de Tim Berners-Lee del CERN como un sistema hipertexto con el único objetivo de servir como medio de transmisión de información entre los científicos que se ocupaban de la Física de alta energía, como parte de la iniciativa *World Wide Web*. Así pues, HTML tuvo lugar a la par que el origen de la Web, ya que se trata del lenguaje que sirve para crear páginas Web.

En 1993 Dan Connolly escribe la primera DTD (*Document Type Definition*) de SGML describiendo el lenguaje y, desde entonces, el lenguaje HTML ha estado sometido a incesantes cambios. De hecho, han existido distintas versiones: 1.0 (en 1993), 2.0 (en 1995), 3.0 (en 1995), 3.2 (en 1997), 4.0 (en 1997, revisada en 1998).

#### **2.4.5 XML**

XML son las siglas del Lenguaje de Etiquetado Extensible. La expresión se forma a partir del acrónimo de la expresión inglesa *eXtensible Markup Language*. Se trata también de un lenguaje estándar que posee una recomendación del *World Wide Web Consortium: Extensible Markup Languages (XML)* [10]. Con la palabra *Extensible* se alude a la no limitación en el número de etiquetas, ya que permite crear aquellas que sean necesarias.

XML surgió como un lenguaje de marcado para sustituir a HTML. Ambos lenguajes son herederos de SGML, el lenguaje de marcas estándar para la descripción formal y de contenido de los documentos, no solamente para la presentación de dichos documentos. El desarrollo de XML comenzó en 1996 y desde entonces ha tenido un desarrollo exponencial. En realidad, XML surge del campo empresarial, ya que HTML era un

Página 75 de 237

lenguaje poco potente para soportar de un modo eficaz y masivo hacer negocios virtuales. Intentando mejorar HTML y tomando como punto de partida el viejo lenguaje SGML, pero simplificándolo para poder trabajar en la Web, se creó XML y sólo 2 años después, en febrero de 1998, fue adoptado como recomendación por el W3C, quien lanzó la versión 1.0. HTML tiene su propia especificación basada en XML, la del lenguaje XHTML (*eXtensible Hypertext Markup Language*) que es, en realidad, un paso intermedio de la migración de HTML hacia XML.

La primera definición de XML fue la de "Sistema para definir, validar y compartir formatos de documentos en la Web". Para crear XML se tomaron las mejores partes tanto del lenguaje SGML como del HTML. La diferencia fundamental entre HTML y XML es que el primero estaba orientado a la presentación de datos, mientras que XML está orientado a los datos en sí mismos, por lo que cualquier software informático trabajará mejor con XML. Sin duda, esta diferencia es fundamental para los nuevos desarrollos de la Web donde se da suma importancia al contenido de los datos y su tratamiento, y no sólo a su presentación.

HTML era, principalmente, un lenguaje de presentación que definía un conjunto de etiquetas y atributos válidos y que ofrecía un significado visual para cada elemento del lenguaje, por el contrario, XML no define las etiquetas ni cómo se utilizan, sino que ofrece un escaso número de reglas sintácticas para poder crear documentos. Así pues, XML no es un lenguaje, sino un metalenguaje o lenguaje para definir otros lenguajes. XML no sustituye a HTML puesto que sirven para cosas distintas: una cosa es presentar la información (para lo que sigue siendo válido HTML) y otra bien distinta es representar e intercambiar los datos de forma independiente a su presentación (que es para lo que sirve XML). Además, XML no sólo se aplica en Internet, sino que se propone como un lenguaje de bajo nivel para intercambio de información estructurada entre distintas plataformas. Se puede utilizar en bases de datos, hojas de cálculo, editores de texto, etc. y no sólo en la Web. HTML y XML son, pues, dos lenguajes complementarios.

Las diferencias fundamentales de XML con respecto a HTML son las siguientes:

- No requiere DTD (*Document Type Definition*).
- Tiene punteros a la estructura de los datos, lo que ahorra tiempo y simplifica el software de aplicación.
- No dispone de soporte para excepciones, por lo que cada etiqueta realiza siempre la misma función.
- Posee independencia de los navegadores y del sistema de objetos, porque en lugar de añadir etiquetas de presentación al documento se remite a una hoja de estilo realizada en XSL (*Extensible Style Language*).

#### 2.4.6 JAVASCRIPT

JavaScript es un lenguaje de *scripting* basado en Objetos, utilizado para acceder a objetos en aplicaciones. Principalmente, se utiliza integrado en un navegador Web permitiendo el desarrollo de interfaces de usuario mejoradas y de páginas Web dinámicas. JavaScript es un dialecto de ECMAScript y se caracteriza por ser un lenguaje basado en prototipos con tipado débil y dinámico con funciones de primera clase. JavaScript ha tenido influencia de múltiples lenguajes y fue diseñado para tener una sintaxis similar a Java, aunque más fácil de utilizar para programadores principiantes.

Todos los navegadores modernos interpretan el código JavaScript integrado dentro de las páginas Web. Para interactuar con una página Web se provee al lenguaje JavaScript de una implementación del DOM.

El lenguaje fue inventado por Brendan Eich en la empresa Netscape Communications, que es la que desarrolló los primeros navegadores Web comerciales. Apareció por primera vez en el producto de Netscape llamado Netscape Navigator 2.0.

Tradicionalmente, se venía utilizando en páginas Web HTML, para realizar tareas y operaciones en el marco de la aplicación únicamente cliente, sin acceso a funciones del servidor. JavaScript se ejecuta en el agente de usuario al mismo tiempo que las sentencias van descargándose junto con el código HTML.

Los autores inicialmente lo llamaron Mocha y más tarde LiveScript pero fue rebautizado como JavaScript en un anuncio conjunto entre Sun Microsystems y Netscape, el 4 de diciembre de 1995.

### **3 ESTUDIO DE LA VIABILIDAD Y PLANTEAMIENTO DEL PROYECTO**

#### **3.1 DESCRIPCIÓN GENERAL DEL SISTEMA**

Dentro de los objetivos del proyecto está el crear una herramienta de escritorio en la que se encuentren integradas las distintas funcionalidades ya introducidas en el Capítulo 1. Por un lado se permitirá extraer la información de unas páginas predefinidas cuyo contenido son eventos de conferencias científicas. Para mayor flexibilidad estas páginas se incluirán en un fichero de configuración. El sistema lee este fichero y muestra los tipos de páginas disponibles para que el usuario inicie el proceso cuando lo desee. La información obtenida sobre los eventos de conferencias científicas, se almacenará en un fichero .xml. En todo momento se irá mostrando el proceso de extracción de información.

Por otro lado, se encuentra la parte de gestión y representación de la información extraída. Para ello se ha de tener en consideración la conexión con las APIs de Google. Para realizar esto será necesario que el usuario tenga creada una cuenta en el gestor de correo de Google: gmail. Esta cuenta permitirá utilizar el servicio de Google Calendar, para crear un nuevo calendario e insertar en él los eventos recuperados en la fase de extracción de información.

En el momento de conectar con Google, el sistema recuperará y mostrará los datos de los eventos que se hayan subido previamente al calendario creado dentro de Google Calendar. Para gestionar esta información, la aplicación local incluirá su propio calendario.

Las localizaciones geográficas donde tendrán lugar los eventos recuperados, se ubicarán en un mapa, utilizando para ello la funcionalidad de representación de Google Maps.

El fichero .xml con los eventos, se podrá almacenar en el disco duro local, con la posibilidad de recuperarlo y de poder subir los eventos en un momento posterior, si por algún motivo no se pudiera conectar con Google, o si no se hubiese deseado subirlos anteriormente a Google Calendar.

La herramienta también permitirá realizar búsquedas de páginas Web y mostrarlas, con la finalidad de poder recuperar manualmente eventos de otras páginas de conferencias científicas. Para ello se permite copiar la información, y pegarla en un formulario o insertarlo a mano. Cuando se guarda el evento, se actualizará en el sistema y se subirá también al control Google Calendar.

### **3.2 DESCRIPCIÓN DEL ENTORNO, USUARIOS Y USOS.**

Esta herramienta va dirigida al personal investigador interesado en organizar y gestionar sus eventos de conferencias científicas. Facilitando por un lado la extracción automática de los eventos de algunas páginas seleccionadas previamente, y la inserción manual de otros eventos de otras páginas.

Los eventos insertados en el sistema se sincronizarán con un calendario creado dentro de Google Calendar para una cuenta de usuario de Google (gmail). El usuario también debe estar familiarizado con el entorno de Google Calendar, si desea poder ver los eventos desde la Web.

En la aplicación Web de Google Calendar los eventos están ubicados en un calendario, permitiendo visualizar el título de los eventos para cada día del calendario. La vista de los días del calendario será configurable por el usuario. Para acceder a la información asociada con cada evento, hay que seleccionar el evento concreto ubicado en el día del calendario correspondiente. Esto implica que sólo se puede acceder a la información de un evento cada vez. La aplicación de escritorio, sin embargo, además de la funcionalidad de ubicar eventos en un calendario, permitirá visualizar los datos de todos los eventos a la vez a través una lista. Esta lista será ordenable por todos sus campos, lo que facilitará su consulta. También se dispondrán de filtros de selección de eventos, facilitando la capacidad de organización y la flexibilidad del sistema de consulta.



### 3.3 ALCANCE DEL PROYECTO

A continuación se mostrará el catálogo de requisitos software encargados de cubrir la funcionalidad requerida para el sistema.

Por una parte, en la definición de requisitos de usuario se recoge qué es lo que el cliente quiere que realice la aplicación, identificando las funcionalidades y restricciones. Existen dos tipos de requisitos de usuario: de capacidad y de restricción.

A continuación de los requisitos de usuario, se presentan los requisitos de sistema, que son una referencia para verificar el diseño y el producto, pues ambos deben cubrir los requisitos que se hayan marcado. Se dividen en funcionales, de rendimiento, de interfaz, de operación, de recursos, de comprobación, de aceptación, de documentación, de seguridad y de calidad.

El campo ID hace referencia al identificador del requisito, que debe ser unívoco. El formato a seguir es el siguiente: R<tipo>-<números>. En este caso el campo <tipo> se corresponde con caracteres del nombre de tipo de requisito (los cuales se definen en la Tabla 23) y <números> son números enteros consecutivos incrementados en una unidad.

TIPO DE REQUISITO	<tipo>
Capacidad	CP
Restricción	RS
Funcional	FN
Rendimiento	RD
Interfaz	IN
Operación	OP
Recursos	RC
Comprobación	CM
Documentación	DC
Seguridad	SG

**Tabla 23.** Definición de identificadores de tipos de requisitos

La especificación de requisitos se ha separado en dos grupos, por un lado se han agrupado los requisitos de usuario, por el otro, los requisitos del sistema.

### 3.3.1 REQUISITOS DE USUARIO

La siguiente tabla (Tabla 24) contiene los requisitos de usuario identificados:

ID	Nombre
RCP-001	Cargar y guardar datos desde un fichero de configuración.
RCP-002	Acceder al contenido de algunas páginas Web de eventos de conferencias científicas e interpretarlo, gestionando los eventos.
RCP-003	Acceder a Google Calendar a través de una cuenta de usuario de Google.
RCP-004	Crear un calendario en Google Calendar para el usuario especificado.
RCP-005	Sincronizar los eventos recuperados o generados en el sistema, con el calendario creado en Google Calendar.
RCP-006	Modularidad del sistema para incluir nuevas páginas de conferencias de las que extraer información automáticamente
RRS-001	El usuario debe tener creada una cuenta de usuario de Google (gmail).
RRS-002	Las páginas que se interpretarán están predefinidas en el sistema.

**Tabla 24.** Especificación de requisitos de usuario

### 3.3.2 REQUISITOS DEL SISTEMA

La siguiente tabla (Tabla 25) contiene los requisitos del sistema:

ID	Nombre
RFN-001	Extraer automáticamente información de eventos de algunas páginas concretas de conferencias científicas.
RFN -002	Permitir guardar los datos de los eventos obtenidos en un fichero local.
RFN -003	Permitir recuperar los datos de los eventos de un fichero local guardado previamente.
RFN -004	Insertar los eventos recuperados dentro de un calendario en Google Calendar para una cuenta de usuario.
RFN -005	Recuperar datos de los eventos que se hayan insertado en Google Calendar.
RFN -006	Mostrar la información de los eventos en una lista.
RFN -007	Mostrar dinámicamente las localizaciones de los eventos en un mapa de Google Maps.
RFN -008	Modificar datos de un evento seleccionado en la lista de eventos.
RFN -009	Eliminar el evento seleccionado en la lista de eventos, pidiendo confirmación previamente.
RFN -010	Eliminar todos los eventos de Google Calendar, pidiendo confirmación previamente.
RFN -011	Permitir ver todos los eventos en la lista de eventos, sin filtros.
RFN -012	Permitir ver los eventos filtrados para un día concreto en la lista de eventos.
RFN -013	Permitir ver los eventos filtrados por nombre en la lista de eventos.
RFN -014	Permitir ver los eventos filtrados por localización en la lista de eventos.
RFN -015	Permitir ver los eventos filtrados por un rango de fechas en la lista de eventos.
RFN -016	Incluir un buscador de páginas Web utilizando Google Search.
RFN -017	Inserción manual de eventos en Google Calendar.
RFN -018	Guardar datos de configuración en fichero de configuración.
RRD-001	El sistema debe funcionar para cualquier usuario que tenga una cuenta de Google, cambiando los datos de acceso en tiempo de ejecución.
RRD-002	El sistema debe sincronizar los eventos de la aplicación de escritorio con los del calendario que crea en Google Calendar en tiempo de ejecución.
RIN-001	Mostrar de forma independiente las distintas partes del sistema: Extracción de

ID	Nombre
	información, Obtener y mostrar eventos de Google Calendar, Búsqueda e Inserción Manual de Eventos y Configuración.
RIN-002	Pestaña de extracción de información y subida de eventos obtenidos a Google Calendar.
RIN-003	Pestaña de recuperación de eventos de Google Calendar y mostrar información de eventos.
RIN-004	Pestaña de búsqueda e inserción manual de eventos en Google Calendar.
RIN-005	Pestaña de datos de configuración.
RIN-006	Opciones de selección de las páginas disponibles para extraer información.
RIN-007	Botón de extracción de información.
RIN-008	Botón para guardar el documento XML en fichero físico local.
RIN-009	Botón para recuperar el documento XML de fichero físico local.
RIN-010	Botón para almacenar los eventos recuperados en Google Calendar.
RIN-011	Control para mostrar los mensajes resultantes de la extracción de información.
RIN-012	Control para mostrar el documento XML con los eventos obtenidos de la extracción.
RIN-013	Botón para recuperar los eventos de Google Calendar.
RIN-014	Opciones de selección de filtros para: Mostrar todos los eventos, mostrar eventos para un día concreto, mostrar eventos por localización, mostrar eventos por nombre y mostrar eventos por rango de fechas.
RIN-015	Control de lista para mostrar los eventos.
RIN-016	Botón para permitir la eliminación de todos los eventos de Google Calendar.
RIN-017	Mostrar menú contextual con opciones de modificar o eliminar sobre el elemento seleccionado en el control de lista de eventos.
RIN-018	Mapa de Google Maps para mostrar las ubicaciones de los eventos.
RIN-019	Facilitar formulario de edición de un evento existente en Google.
RIN-020	Facilitar formulario de inserción manual de un nuevo evento.
RIN-021	Control de búsqueda de páginas Web de Google Search.
RIN-022	Controles para mostrar y modificar los datos del fichero de configuración agrupados de la siguiente forma: Cuenta de usuario de Google, Fechas de recuperación de eventos, opciones de configuración de proxy, nombre de calendario a crear en Google.
ROP-001	Inicialización de listas de eventos y documento XML para los eventos al iniciar la aplicación.

ID	Nombre
ROP-002	Recuperar datos de la cuenta de usuario del fichero de configuración.
ROP-003	Descarga de contenido de páginas Web.
ROP-004	Interpretación del contenido descargado en función del tipo de página.
ROP-005	Interpretación y tratamiento de los eventos en función de las fechas que contengan los eventos recuperados de las páginas, distinguiendo así los siguientes tipos de eventos: <main conference>, <deadline>, <deadline abstract>, <deadline full submission> y <deadline workshop>.
ROP-006	Almacenar los eventos recuperados en un documento XML.
ROP-007	Conectar con el servicio de Google Calendar.
ROP-008	Insertar los eventos recuperados en un calendario de Google Calendar.
ROP-009	Seleccionar una ubicación al guardar un documento XML en un fichero físico.
ROP-010	Recuperar de una ubicación el fichero con los datos de los eventos.
ROP-011	Validar el documento XML generado, contra un esquema de validación. <sup>15</sup>
ROP-012	Insertar mensajes de error en un fichero de LOG.
ROP-013	Almacenar los eventos recuperados de Google Calendar en una estructura en memoria.
ROP-014	Gestionar las cargas de los eventos en el control de lista, y la carga de las ubicaciones en el mapa a través de la estructura interna.
ROP-015	Actualizar los controles afectados después de eliminar o modificar cualquier evento.
ROP-016	Vaciar todos los controles del interfaz, variables y estructuras en memoria cuando se eliminan todos los eventos.
ROP-017	Vaciar todos los controles del interfaz, variables y estructuras en memoria cuando se cambia el nombre del calendario.
ROP-018	Vaciar todos los controles del interfaz, variables y estructuras en memoria cuando se cambian los datos de acceso de la cuenta de gmail de un usuario.
ROP-019	Mostrar información sobre el estado del proceso de extracción de información.
ROP-020	Mostrar un documento XML con los datos de los eventos recuperados del proceso de extracción de información.
ROP-021	Mostrar ruta del fichero físico guardado y/o recuperado, que contiene los eventos.
ROP-022	Selección de los distintos filtros disponibles sobre los eventos recuperados de Google Calendar.
ROP-023	Carga automática de eventos en el control lista y las ubicaciones en el mapa al recuperar los eventos de Google Calendar.

ID	Nombre
ROP-024	Carga automática de eventos en el control lista y las ubicaciones en el mapa al seleccionar cualquier filtro.
ROP-025	Carga automática de eventos en el control lista y las ubicaciones en el mapa cuando se filtra por nombre, al insertar un carácter.
ROP-026	Tratar y normalizar los formatos de fecha extraídos de las diferentes fuentes de datos.
RRC-001	Disponer de una conexión a Internet con un ancho de banda adecuado para acceder a las páginas Web al extraer la información y también para los accesos a Google Calendar.
RCM-001	Comprobar antes de extraer información automática, que se ha seleccionado alguna página.
RCM-002	Comprobaciones de datos al insertar un nuevo evento y al modificar un evento existente.
RCM-003	Comprobar la conexión con el servicio de Google Calendar, para el usuario especificado, antes de realizar cualquier operación que necesite acceder a Google, mostrando aviso cuando no pueda conectar.
RCM-004	Comprobar la aceptación del mensaje de aviso de un usuario antes de eliminar eventos en Google Calendar.
RCM-005	Comprobar que no se inserten eventos duplicados en Google Calendar. Para ello se comprueba, el nombre del evento, la ubicación y la fecha asociada.
RCM-006	Comprobaciones de datos cuando se cambian los datos de configuración, avisando de lo que implica al usuario.
RDC-001	Manual de usuario.
RDC-002	Manual de instalación.
RDC-003	Comentarios del código fuente.
RDC-004	Tipo y tamaño de letra en la documentación.
RDC-005	Legibilidad del manual de usuario.
RDC-006	Idioma español.
RSG-001	Mostrar la contraseña del usuario en la aplicación con máscara de caracteres “*”.
RSG-002	Aplicar un algoritmo de codificación sobre la contraseña de la cuenta de usuario, cuando se vaya a insertar en el fichero de configuración, de manera que si se edita ese fichero dicha contraseña no sea legible.

**Tabla 25.** Especificación de requisitos del sistema

### **3.4 ANÁLISIS DE VIABILIDAD**

#### **3.4.1 SOFTWARE**

A continuación se va a evaluar el software necesario para cumplir con los requisitos especificados en el punto anterior.

El proyecto se llevará a cabo usando el sistema operativo Windows como plataforma de ejecución. En principio se utilizará Windows Vista. También será compatible con otras plataformas Windows anteriores, como Windows XP y Windows Server 2003.

Sobre el entorno de ejecución, se ha estudiado la posibilidad de utilizar Java o .NET 2005 (concretamente C#). Se ha elegido .NET porque en la documentación del API de Datos de Google Calendar se especifica que las librerías cliente del API de Datos de Google Calendar para .NET, permiten integrar Google Calendar en aplicaciones de escritorio. De las librerías cliente del entorno Java sólo se comenta que están preparadas para usarse en aplicaciones Web. Se descarta por tanto, el entorno Java.

Para la fase de extracción de información, se plantea la utilización de expresiones regulares para el parseo de código fuente que se recuperará de las páginas Web para facilitar su creación y pruebas, se utilizará la herramienta Regulator 2.03.

Para la comprobación del funcionamiento de la aplicación cuando se hagan las cargas de eventos a Google calendar, se necesitará utilizar un navegador. Para ello, se puede utilizar cualquier navegador que acceda sin problema al gestor de contenidos de Google: gmail. Se proponen, por ejemplo, los dos siguientes: Internet Explorer 8.0 y Firefox 3.5.5. Se ha comprobado que con cualquiera de los dos se accede perfectamente.



### **3.4.2 HARDWARE**

Será necesario disponer de una máquina capaz de procesar las páginas en un tiempo razonable y con conexión a Internet con un ancho de banda suficientemente amplio para que los accesos a las páginas no requieran demasiado tiempo.

Por tanto, los requisitos hardware que se proponen son los siguientes:

- Procesador Intel® Core™ 2 Duo CPU T5750 2.00 GHz.
- Tarjeta Gráfica NVIDIA GeForce 8400M GS.
- 4Gb de memoria RAM.
- Disco duro de 200Gb.

### 3.5 PRUEBAS DE VALIDACIÓN

En este punto se van a enumerar las pruebas necesarias para comprobar que el sistema cumple con todos los requisitos marcados por el usuario (Tabla 26). Son necesarias para que el desarrollador valide el sistema antes de la puesta en explotación.

Se debe insistir sobre todo en probar que el sistema cumple los criterios de aceptación del sistema que sirven de base para asegurar que satisface los requisitos exigidos. Estos criterios de aceptación quedan definidos de forma clara prestando atención a: procesos identificados como críticos, rendimiento, seguridad y disponibilidad del sistema. Los puntos de atención dependen de la naturaleza del sistema y los requisitos marcados por el usuario.

ID	Nombre
PR-001	Obtener eventos de una página Web
PR-002	Comprobación de eventos obtenidos
PR-003	Guardar el documento XML con los eventos en disco duro
PR-004	Recuperar el fichero con los eventos del disco duro
PR-005	Inserción de eventos en Google Calendar
PR-006	Recuperar y mostrar eventos de Google Calendar
PR-007	Pruebas de filtros de eventos del sistema
PR-008	Modificar los datos de un evento de la lista de eventos
PR-009	Eliminar un evento de la lista de eventos
PR-010	Eliminar todos los eventos del sistema
PR-011	Buscar páginas con el buscador de Google Search
PR-012	Añadir evento de forma manual
PR-013	Cambiar los datos de acceso del usuario en pestaña configuración
PR-014	Cambiar fechas de recuperación de eventos en pestaña configuración
PR-015	Probar opciones de proxy en pestaña configuración
PR-016	Introducir nombre de Calendario en pestaña configuración

**Tabla 26.** Descripción de las prueba del sistema

A continuación se describen las pruebas con detalle:

PR-001	
Nombre	Obtener eventos de una página Web
Descripción	Obtener los eventos de una página de conferencias científicas de forma automática
Elementos relacionados	RCP-002, RFN-001
Especificaciones de entrada	Ninguna
Especificaciones de salida	Resultados del proceso mostrado en pantalla, documento XML con todos los eventos recuperados
Entorno de prueba	Conexión a Internet y servidores de alojamiento de las páginas disponibles

**Tabla 27.** Descripción de la prueba PR-001

PR-002	
Nombre	Comprobación de eventos obtenidos
Descripción	Comprobar que los eventos obtenidos, que se encuentran en el documento XML, se corresponden con los eventos que se encuentran en la página Web
Elementos relacionados	RCP-002, RFN -001
Especificaciones de entrada	Documento XML con los eventos
Especificaciones de salida	Ninguna
Entorno de prueba	Conexión a Internet y servidores de alojamiento de las páginas disponibles

**Tabla 28.** Descripción de la prueba PR-002

PR-003	
Nombre	Guardar el documento XML con los eventos en disco duro
Descripción	Guardar el documento XML en un fichero físico. Comprobar que el fichero se ha guardado correctamente
Elementos relacionados	RFN -002
Especificaciones de entrada	Documento XML con los eventos
Especificaciones de salida	Fichero creado en la ruta especificada del disco duro
Entorno de prueba	Ninguno

**Tabla 29.** Descripción de la prueba PR-003

PR-004	
Nombre	Recuperar el fichero con los eventos del disco duro
Descripción	Recuperar los datos de los eventos de un fichero físico. Comprobar que los eventos se han recuperado correctamente en el documento XML
Elementos relacionados	RFN -003
Especificaciones de entrada	Fichero con los eventos en el disco duro
Especificaciones de salida	Documento XML en el sistema
Entorno de prueba	Ninguno

**Tabla 30.** Descripción de la prueba PR-004

PR-005	
Nombre	Insertión de eventos en Google Calendar
Descripción	Insertar los eventos recuperados del proceso de extracción de información en Google Calendar, creando para ello un nuevo calendario. Acceder a través de un navegador al gestor de correo de Google, con el mismo usuario y contraseña con el que se accedió a la aplicación, y comprobar que el calendario se ha creado correctamente y que contiene los eventos insertados, con sus datos correctos
Elementos relacionados	RCP-003, RCP-004, RCP-005, RFN -004, RRD-002
Especificaciones de entrada	Eventos recuperados del proceso de extracción de información
Especificaciones de salida	Calendario creado con Eventos insertados dentro de Google Calendar
Entorno de prueba	Conexión a Internet y servidores de Google disponibles

**Tabla 31.** Descripción de la prueba PR-005

PR-006	
Nombre	Recuperar y mostrar eventos de Google Calendar
Descripción	Recuperar los eventos del calendario creado en Google Calendar. Comprobar que los eventos que se muestran en la lista de eventos de la aplicación son los todos los que se encuentran en Google Calendar
Elementos relacionados	RCP-003, RCP-005, RFN -005, RFN -006, RFN -007, RRD-002
Especificaciones de entrada	Ninguna
Especificaciones de salida	Los eventos mostrados en la lista de eventos del sistema, y sus ubicaciones en un mapa.
Entorno de prueba	Conexión a Internet y servidores de Google disponibles

**Tabla 32.** Descripción de la prueba PR-006

PR-007	
Nombre	Pruebas de filtros de eventos del sistema
Descripción	Probar los filtros de eventos comprobando que los eventos mostrados en la lista son los correspondientes a cada filtro
Elementos relacionados	RFN -011, RFN -012, RFN -013, RFN -014, RFN -015
Especificaciones de entrada	Eventos recuperados de Google Calendar
Especificaciones de salida	Eventos mostrados en la lista con sus ubicaciones en el mapa, para cada filtro
Entorno de prueba	Ninguno

**Tabla 33.** Descripción de la prueba PR-007

PR-008	
Nombre	Modificar los datos de un evento de la lista de eventos
Descripción	Modificar un evento seleccionado en la lista, comprobando que todos los datos se actualizan correctamente. Y que todos los controles se refrescan con los nuevos datos, incluida la ubicación del evento en el filtro de localizaciones y en el mapa, en ambos casos, si procede.
Elementos relacionados	RCP-003, RFN -008
Especificaciones de entrada	Lista de eventos con eventos cargados
Especificaciones de salida	Eventos mostrados en la lista con sus ubicaciones en el mapa
Entorno de prueba	Conexión a Internet y servidores de Google disponibles

**Tabla 34.** Descripción de la prueba PR-008

PR-009	
Nombre	Eliminar un evento de la lista de eventos
Descripción	Eliminar un evento seleccionándolo en la lista de eventos, comprobando posteriormente que el evento se elimina de la lista de eventos, así como, la ubicación del evento del filtro de localizaciones y del mapa, si procede.
Elementos relacionados	RCP-003, RFN-009
Especificaciones de entrada	Lista de eventos con eventos cargados
Especificaciones de salida	Eventos mostrados en la lista con sus ubicaciones en el mapa
Entorno de prueba	Conexión a Internet y servidores de Google disponibles

**Tabla 35.** Descripción de la prueba PR-009

PR-010	
Nombre	Eliminar todos los eventos del sistema
Descripción	Eliminar todos los eventos. Comprobando que se eliminan del calendario de Google Calendar y que se vacían la lista del filtro de localizaciones y el mapa con las ubicaciones.
Elementos relacionados	RCP-003, RFN-010
Especificaciones de entrada	Lista de eventos con eventos cargados
Especificaciones de salida	Todos los controles asociados vacíos
Entorno de prueba	Conexión a Internet y servidores de Google disponibles

**Tabla 36.** Descripción de la prueba PR-010

PR-011	
Nombre	Buscar páginas con el buscador de Google Search
Descripción	Comprobar que el buscador de páginas Web funciona correctamente.
Elementos relacionados	RFN-016
Especificaciones de entrada	Texto introducido
Especificaciones de salida	Páginas encontradas
Entorno de prueba	Conexión a Internet y servidores de Google disponibles

**Tabla 37.** Descripción de la prueba PR-011

PR-012	
Nombre	Añadir evento de forma manual
Descripción	Añadir un nuevo evento al calendario de Google Calendar de forma manual. Observar que se añade correctamente en la aplicación y en Google Calendar. Refrescar los eventos recuperados para comprobar que el sistema lo recupera correctamente.
Elementos relacionados	RCP-003, RFN-017
Especificaciones de entrada	Datos del nuevo evento
Especificaciones de salida	Evento insertado en el calendario de Google Calendar
Entorno de prueba	Conexión a Internet y servidores de Google disponibles

**Tabla 38.** Descripción de la prueba PR-012



PR-013	
Nombre	Cambiar los datos de acceso del usuario en pestaña de configuración
Descripción	Insertar la cuenta de otro usuario de Google en la pestaña de configuración, guardándolo en el fichero de configuración. Comprobar que se vacían los controles y posteriormente, recuperar los eventos del nuevo usuario, si los tuviese.
Elementos relacionados	RCP-001, RFN-018, RRD-001
Especificaciones de entrada	Datos del nuevo usuario
Especificaciones de salida	Eventos recuperados para el nuevo usuario
Entorno de prueba	Conexión a Internet y servidores de Google disponibles

**Tabla 39.** Descripción de la prueba PR-013

PR-014	
Nombre	Cambiar fechas de recuperación de eventos en pestaña de configuración
Descripción	Cambiar las fechas de recuperación de eventos de la pestaña de configuración, comprobando que efectivamente, todos los eventos que se recuperan del calendario de Google Calendar, son los que se encuentran en ese rango de fechas.
Elementos relacionados	RFN-018
Especificaciones de entrada	Ninguna
Especificaciones de salida	Eventos recuperados para el nuevo rango de fechas
Entorno de prueba	Conexión a Internet y servidores de Google disponibles

**Tabla 40.** Descripción de la prueba PR-014

PR-015	
Nombre	Probar opciones de proxy en pestaña de configuración
Descripción	Si el sistema requiere un proxy para conectarse a Internet, introducir el servidor y el puerto, desmarcar la opción de “No utilizar proxy”. Si por el contrario, no necesita proxy, marcar la opción anterior. Y en cualquiera de los dos casos, probar que el sistema funciona.
Elementos relacionados	RFN-018
Especificaciones de entrada	Datos del proxy o marca en opción de “No utilizar proxy”
Especificaciones de salida	Acceso a Internet con proxy, o sin él
Entorno de prueba	Conexión a Internet y servidores de Google disponibles

**Tabla 41.** Descripción de la prueba PR-015

PR-016	
Nombre	Introducir nombre de Calendario en pestaña de configuración
Descripción	Introducir un nombre de calendario en la pestaña de configuración y hacer clic en guardar. Comprobar que se eliminan los datos cargados en el sistema, si los hubiera. Comprobar que las operaciones de acceso a Google, se realizan creando el nuevo calendario.
Elementos relacionados	RFN-018
Especificaciones de entrada	Nombre de Calendario
Especificaciones de salida	Calendario creado en Google Calendar
Entorno de prueba	Conexión a Internet y servidores de Google disponibles

**Tabla 42.** Descripción de la prueba PR-016

### 3.6 METODOLOGÍA DEL DESARROLLO

Para la realización de este proyecto se ha utilizado un ciclo de vida orientado a objetos.

Los objetos, además de centrarse en el análisis y en el diseño, como otros ciclos de vida del software, tienen una particularidad, y es que están basados en componentes que se relacionan entre ellos a través de interfaces, o lo que es lo mismo, son más modulares y por lo tanto el trabajo se puede dividir en un conjunto de miniproyectos. Además, hoy en día la tendencia es a reducir los riesgos, y en este sentido, el ciclo de vida en cascada no proporciona muchas facilidades. Debido a todo esto, el ciclo de vida típico en una metodología de diseño orientado a objetos es iterativo e incremental.

A continuación se muestra el esquema del tipo de ciclo de vida orientado a objetos aplicado: el modelo fuente.

#### **Modelo fuente [9]**

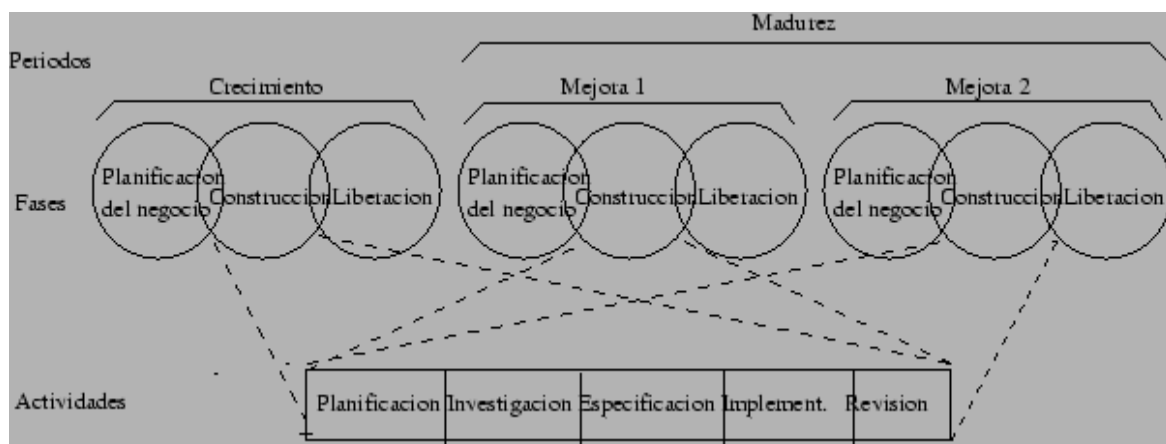
Fue creado por Henderson-Sellers y Edwards en 1990. Es un tipo de ciclo de vida pensado para la orientación a objetos y posiblemente el más seguido. Un proyecto se divide en las fases:

- Planificación del negocio
- Construcción: Es la más importante y se divide a su vez en otras cinco actividades:
  - Planificación
  - Investigación
  - Especificación
  - Implementación
  - Revisión
- Entrega

La primera y la tercera fase son independientes de la metodología de desarrollo orientado a objetos. Además de las tres fases, existen dos periodos:

- Crecimiento: Es el tiempo durante el cual se construye el sistema
- Madurez: Es el periodo de mantenimiento del producto. Cada mejora se planifica igual que el periodo anterior, es decir, con las fases de Planificación del negocio, Construcción y Entrega.

Cada clase puede tener un ciclo de vida sólo para ella debido a que cada una puede estar en una fase diferente en un momento cualquiera. La ventaja es que permite un desarrollo solapado e iterativo. En la Figura 17 se muestra un esquema de este tipo de ciclo de vida.



**Figura 17.** Ciclo de vida fuente

Durante el ciclo de vida del proyecto se hizo una primera planificación con el cliente, con una amplia especificación de requisitos, intentando definir lo que se podía esperar del sistema. Algunos de estos requisitos, no se sabía previamente si eran viables o no, requerían de una investigación profunda para saber si se podían aplicar o si había que replantearlo y adaptarlo de otras formas.

En la fase de construcción, el sistema ha pasado por distintas etapas, como este modelo indica:

Se ha hecho una planificación inicial, ha tenido una etapa de investigación, otra de especificación, pasando después a la implementación, con un primer prototipo de cómo se podía plantear el sistema, que posteriormente fue revisado por parte del cliente.

Después de estas revisiones con el cliente, han surgido nuevas ideas y posibles ampliaciones y mejoras del sistema, pasando de nuevo por todo el ciclo de la fase de construcción. Este ciclo se ha repetido varias veces, hasta conseguir lo que la aplicación es hoy en día.

El sistema está orientado a objetos en su totalidad y se ha modularizado como más adelante se detallará en el apartado dedicado a la implementación. La funcionalidad agrupada en cada módulo es totalmente independiente de los demás módulos.

## 4 ANÁLISIS Y DISEÑO DEL SISTEMA

El sistema se puede separar cuatro partes independientes entre ellas:

1. La extracción automática de información. En esta parte se realizará la obtención de eventos de las páginas Web de conferencias científicas predefinidas en el sistema. Los eventos obtenidos de este proceso se deben poder insertar dentro de un calendario en Google Calendar. Este calendario, será creado previamente por el sistema dentro de Google Calendar. El nombre de dicho calendario podrá ser establecido por el usuario.
2. La obtención y gestión de eventos del calendario creado previamente en Google Calendar. En esta parte, el sistema debe mostrar en una lista los eventos obtenidos, y permitir modificar sus datos y eliminarlos del calendario de Google Calendar. Las ubicaciones de los eventos que se estén mostrando en cada momento en la lista deberán aparecer localizadas en un mapa de Google Maps. Cuando se seleccione un marcador dentro del mapa, se deben mostrar la dirección del evento, y los eventos que tengan lugar en esa ubicación.

También se debe permitir que los eventos se muestren en función de varias opciones de filtrado:

- Mostrar todos los eventos.
  - Mostrar los eventos para un día concreto.
  - Mostrar los eventos filtrando el nombre por los caracteres introducidos por el usuario.
  - Mostrar los eventos para una ubicación concreta.
  - Mostrar los eventos que se encuentre entre un rango de fechas.
3. La búsqueda de páginas Web, a través de un buscador de Google Search, y la inserción de nuevos eventos de forma manual, rellenando sus datos en un formulario. Los controles de inserción de texto del formulario deben permitir copiar

y pegar texto. Cuando se guarden los datos, el nuevo evento se insertará en el calendario de eventos creado en Google Calendar.

4. Por último tiene que existir una parte donde se recuperen y se permita la modificación de los datos de configuración. Estos datos, se agrupan en cuatro puntos:

- Datos de acceso a Google para un usuario:
  - Nombre de usuario
  - Contraseña
- Datos de configuración de un *proxy*:
  - Servidor
  - Puerto
- Fechas de recuperación de eventos de Google Calendar:
  - Fecha de Inicio
  - Fecha de Fin
- Nombre del calendario en Google Calendar:
  - Nombre de calendario

## 4.1 ESPECIFICACIÓN DE REQUISITOS

En el apartado 3.3 se mostraba el catálogo de requisitos software encargados de cubrir la funcionalidad requerida para el sistema. A continuación se muestra detalladamente la descripción de cada uno de esos requisitos identificados. Se han agrupado en requisitos de usuario y requisitos de sistema.

Los niveles de **Prioridad**, de **Claridad** y de **Verificabilidad** pueden ser:

- Alta
- Media
- Baja

Los niveles de **Necesidad** pueden ser:

- Esencial
- Deseable
- Opcional

### 4.1.1 REQUISITOS DE USUARIO

RCP-001	
Prioridad	Media
Necesidad	Opcional
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Cargar y guardar datos desde un fichero de configuración.

*Tabla 43.* Descripción del requisito RCP-001



RCP-002	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Acceder al contenido de algunas páginas Web de eventos de conferencias científicas e interpretarlo, gestionando los eventos.

**Tabla 44.** Descripción del requisito RCP-002

RCP-003	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Acceder a Google Calendar a través de una cuenta de usuario de Google.

**Tabla 45.** Descripción del requisito RCP-003

RCP-004	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Crear un calendario en Google Calendar para el usuario especificado.

**Tabla 46.** Descripción del requisito RCP-004

RCP-005	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Sincronizar los eventos recuperados o generados en el sistema, con el calendario creado en Google Calendar.

**Tabla 47.** Descripción del requisito RCP-005

RCP-006	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Modularidad del sistema para incluir nuevas páginas de conferencias de las que extraer información automáticamente.

**Tabla 48.** Descripción del requisito RCP-006

RRS-001	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	El usuario debe tener creada una cuenta de usuario de Google (gmail).

**Tabla 49.** Descripción del requisito RRS-001

RRS-002	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Las páginas que se interpretarán están predefinidas en el sistema.

**Tabla 50.** Descripción del requisito RRS-002

#### 4.1.2 REQUISITOS DE SISTEMA

RFN-001	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Extraer automáticamente información de eventos de algunas páginas concretas de conferencias científicas.

**Tabla 51.** Descripción del requisito RFN-001

RFN-002	
Prioridad	Media
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Permitir guardar los datos de los eventos obtenidos en un fichero local.

**Tabla 52.** Descripción del requisito RFN-002

RFN-003	
Prioridad	Media
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Permitir recuperar los datos de los eventos de un fichero local guardado previamente.

**Tabla 53.** Descripción del requisito RFN-003

RFN-004	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Insertar los eventos recuperados dentro de un calendario en Google Calendar para una cuenta de usuario.

**Tabla 54.** Descripción del requisito RFN-004

RFN-005	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Recuperar datos de los eventos que se hayan insertado en Google Calendar.

**Tabla 55.** Descripción del requisito RFN-005

RFN-006	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Mostrar la información de los eventos en una lista.

**Tabla 56.** Descripción del requisito RFN-006

RFN-007	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Mostrar dinámicamente las localizaciones de los eventos en un mapa de Google Maps.

**Tabla 57.** Descripción del requisito RFN-007

RFN-008	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Modificar datos de un evento seleccionado en la lista de eventos.

**Tabla 58.** Descripción del requisito RFN-008

RFN-009	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Eliminar el evento seleccionado en la lista de eventos, pidiendo confirmación previamente.

**Tabla 59.** Descripción del requisito RFN-009



RFN-010	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Eliminar todos los eventos de Google Calendar, pidiendo confirmación previamente.

**Tabla 60.** Descripción del requisito RFN-010

RFN-011	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Permitir ver todos los eventos en la lista de eventos, sin filtros.

**Tabla 61.** Descripción del requisito RFN-011

RFN-012	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Permitir ver los eventos filtrados para un día concreto en la lista de eventos.

**Tabla 62.** Descripción del requisito RFN-012

RFN-013	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Permitir ver los eventos filtrados por nombre en la lista de eventos.

**Tabla 63.** Descripción del requisito RFN-013

RFN-014	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Permitir ver los eventos filtrados por localización en la lista de eventos.

**Tabla 64.** Descripción del requisito RFN-014

RFN-015	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Permitir ver los eventos filtrados por un rango de fechas en la lista de eventos.

**Tabla 65.** Descripción del requisito RFN-015

RFN-016	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Incluir un buscador de páginas Web utilizando Google Search.

**Tabla 66.** Descripción del requisito RFN-016

RFN-017	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Inserción manual de eventos en Google Calendar.

**Tabla 67.** Descripción del requisito RFN-017

RFN-018	
Prioridad	Media
Necesidad	Opcional
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Guardar datos de configuración en fichero de configuración.

**Tabla 68.** Descripción del requisito RFN-018

RRD-001	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	El sistema debe funcionar para cualquier usuario que tenga una cuenta de Google, cambiando los datos de acceso en tiempo de ejecución.

**Tabla 69.** Descripción del requisito RRD-001

RRD-002	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	El sistema debe sincronizar los eventos de la aplicación de escritorio con los del calendario que crea en Google Calendar en tiempo de ejecución.

**Tabla 70.** Descripción del requisito RRD-002

RIN-001	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Mostrar de forma independiente las distintas partes del sistema: Extracción de información, Obtener y mostrar eventos de Google Calendar, Búsqueda e Inserción Manual de Eventos y Configuración.

**Tabla 71.** Descripción del requisito RIN-001

RIN-002	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Pestaña de extracción de información y subida de eventos obtenidos a Google Calendar.

**Tabla 72.** Descripción del requisito RIN-002

RIN-003	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Pestaña de recuperación de eventos de Google Calendar y mostrar información de eventos.

**Tabla 73.** Descripción del requisito RIN-003

RIN-004	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Pestaña de búsqueda e inserción manual de eventos en Google Calendar.

**Tabla 74.** Descripción del requisito RIN-004

RIN-005	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Pestaña de datos de configuración.

**Tabla 75.** Descripción del requisito RIN-005

RIN-006	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Opciones de selección de las páginas disponibles para extraer información.

**Tabla 76.** Descripción del requisito RIN-006

RIN-007	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Botón de extracción de información.

**Tabla 77.** Descripción del requisito RIN-007



RIN-008	
Prioridad	Media
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Botón para guardar el documento XML en fichero físico local.

**Tabla 78.** Descripción del requisito RIN-008

RIN-009	
Prioridad	Media
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Botón para recuperar el documento XML de fichero físico local.

**Tabla 79.** Descripción del requisito RIN-009

RIN-010	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Botón para almacenar los eventos recuperados en Google Calendar.

**Tabla 80.** Descripción del requisito RIN-010

RIN-011	
Prioridad	Media
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Control para mostrar los mensajes resultantes de la extracción de información.

**Tabla 81.** Descripción del requisito RIN-011

RIN-012	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Control para mostrar el documento XML con los eventos obtenidos de la extracción.

**Tabla 82.** Descripción del requisito RIN-012

RIN-013	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Botón para recuperar los eventos de Google Calendar.

**Tabla 83.** Descripción del requisito RIN-013

RIN-014	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Opciones de selección de filtros para: mostrar todos los eventos, mostrar eventos para un día concreto, mostrar eventos por localización, mostrar eventos por nombre y mostrar eventos por rango de fechas.

**Tabla 84.** Descripción del requisito RIN-014

RIN-015	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Control de lista para mostrar los eventos.

**Tabla 85.** Descripción del requisito RIN-015

RIN-016	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Botón para permitir la eliminación de todos los eventos de Google Calendar.

**Tabla 86.** Descripción del requisito RIN-016

RIN-017	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Mostrar menú contextual con opciones de modificar o eliminar sobre el elemento seleccionado en el control de lista de eventos.

**Tabla 87.** Descripción del requisito RIN-017

RIN-018	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Mapa de Google Maps para mostrar las ubicaciones de los eventos.

**Tabla 88.** Descripción del requisito RIN-018

RIN-019	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Facilitar formulario de edición de un evento existente en Google.

**Tabla 89.** Descripción del requisito RIN-019

RIN-020	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Facilitar formulario de inserción manual de un nuevo evento.

**Tabla 90.** Descripción del requisito RIN-020

RIN-021	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Control de búsqueda de páginas Web de Google Search.

**Tabla 91.** Descripción del requisito RIN-021

RIN-022	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Controles para mostrar y modificar los datos del fichero de configuración agrupados de la siguiente forma: Cuenta de usuario de Google, Fechas de recuperación de eventos, opciones de configuración de proxy, nombre de calendario a crear en Google.

**Tabla 92.** Descripción del requisito RIN-022

ROP-001	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Inicialización de listas de eventos y documento XML para los eventos al iniciar la aplicación.

**Tabla 93.** Descripción del requisito ROP-001



ROP-002	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Recuperar datos de la cuenta de usuario del fichero de configuración.

**Tabla 94.** Descripción del requisito ROP-002

ROP-003	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Descarga de contenido de páginas Web.

**Tabla 95.** Descripción del requisito ROP-003

ROP-004	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Interpretación del contenido descargado en función del tipo de página.

**Tabla 96.** Descripción del requisito ROP-004

ROP-005	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Interpretación y tratamiento de los eventos en función de las fechas que contengan los eventos recuperados de las páginas, distinguiendo así los siguientes tipos de eventos: <main conference>, <deadline>, <deadline abstract>, <deadline full submission> y <deadline workshop>.

**Tabla 97.** Descripción del requisito ROP-005

ROP-006	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Almacenar los eventos recuperados en un documento XML.

**Tabla 98.** Descripción del requisito ROP-006

ROP-007	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Conectar con el servicio de Google Calendar.

**Tabla 99.** Descripción del requisito ROP-007

ROP-008	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Insertar los eventos recuperados en un calendario de Google Calendar.

**Tabla 100.** Descripción del requisito ROP-008

ROP-009	
Prioridad	Media
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Seleccionar una ubicación al guardar un documento XML en un fichero físico.

**Tabla 101.** Descripción del requisito ROP-009

ROP-010	
Prioridad	Media
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Recuperar de una ubicación el fichero con los datos de los eventos.

**Tabla 102.** Descripción del requisito ROP-010

ROP-011	
Prioridad	Media
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Validar el documento XML generado, contra un esquema de validación.

**Tabla 103.** Descripción del requisito ROP-011

ROP-012	
Prioridad	Media
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Insertar mensajes de error en un fichero de LOG.

**Tabla 104.** Descripción del requisito ROP-012

ROP-013	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Almacenar los eventos recuperados de Google Calendar en una estructura en memoria.

**Tabla 105.** Descripción del requisito ROP-013

ROP-014	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Gestionar las cargas de los eventos en el control de lista, y la carga de las ubicaciones en el mapa a través de la estructura interna.

**Tabla 106.** Descripción del requisito ROP-014

ROP-015	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Actualizar los controles afectados después de eliminar o modificar cualquier evento.

**Tabla 107.** Descripción del requisito ROP-015

ROP-016	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Vaciar todos los controles del interfaz, variables y estructuras en memoria cuando se eliminan todos los eventos.

**Tabla 108.** Descripción del requisito ROP-016

ROP-017	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Vaciar todos los controles del interfaz, variables y estructuras en memoria cuando se cambia el nombre del calendario.

**Tabla 109.** Descripción del requisito ROP-017

ROP-018	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Vaciar todos los controles del interfaz, variables y estructuras en memoria cuando se cambian los datos de acceso de la cuenta de Google (gmail) de un usuario.

**Tabla 110.** Descripción del requisito ROP-018



ROP-019	
Prioridad	Media
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Mostrar información sobre el estado del proceso de extracción de información.

**Tabla 111.** Descripción del requisito ROP-019

ROP-020	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Mostrar un documento XML con los datos de los eventos recuperados del proceso de extracción de información.

**Tabla 112.** Descripción del requisito ROP-020

ROP-021	
Prioridad	Baja
Necesidad	Opcional
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Mostrar ruta del fichero físico guardado y/o recuperado, que contiene los eventos.

**Tabla 113.** Descripción del requisito ROP-021

ROP-022	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Selección de los distintos filtros disponibles sobre los eventos recuperados de Google Calendar.

**Tabla 114.** Descripción del requisito ROP-022

ROP-023	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Carga automática de eventos en el control lista y las ubicaciones en el mapa al recuperar los eventos de Google Calendar.

**Tabla 115.** Descripción del requisito ROP-023

ROP-024	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Carga automática de eventos en el control lista y las ubicaciones en el mapa al seleccionar cualquier filtro.

**Tabla 116.** Descripción del requisito ROP-024

ROP-025	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Carga automática de eventos en el control lista y las ubicaciones en el mapa cuando se filtra por nombre, al insertar un carácter.

**Tabla 117.** Descripción del requisito ROP-025

ROP-026	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Tratar y normalizar los formatos de fecha extraídos de las diferentes fuentes de datos.

**Tabla 118.** Descripción del requisito ROP-026

RRC-001	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Disponer de una conexión a Internet con un ancho de banda adecuado para acceder a las páginas Web al extraer la información y también para los accesos a Google Calendar.

**Tabla 119.** Descripción del requisito RRC-001

RCM-001	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Comprobar antes de extraer información automática, que se ha seleccionado alguna página.

**Tabla 120.** Descripción del requisito RCM-001

RCM-002	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Comprobaciones de datos al insertar un nuevo evento y al modificar un evento existente.

**Tabla 121.** Descripción del requisito RCM-002

RCM-003	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Comprobar la conexión con el servicio de Google Calendar, para el usuario especificado, antes de realizar cualquier operación que necesite acceder a Google, mostrando aviso cuando no pueda conectar.

**Tabla 122.** Descripción del requisito RCM-003

RCM-004	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Comprobar la aceptación del mensaje de aviso de un usuario antes de eliminar eventos en Google Calendar.

**Tabla 123.** Descripción del requisito RCM-004

RCM-005	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Comprobar que no se inserten eventos duplicados en Google Calendar. Para ello se comprueba, el nombre del evento, la ubicación y la fecha asociada.

**Tabla 124.** Descripción del requisito RCM-005

RCM-006	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Comprobaciones de datos cuando se cambian los datos de configuración, avisando de lo que implica al usuario.

**Tabla 125.** Descripción del requisito RCM-006

RDC-001	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Manual de usuario.

**Tabla 126.** Descripción del requisito RDC-001



RDC-002	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Manual de instalación.

**Tabla 127.** Descripción del requisito RDC-002

RDC-003	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Comentarios del código fuente.

**Tabla 128.** Descripción del requisito RDC-003

RDC-004	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Tipo y tamaño de letra en la documentación.

**Tabla 129.** Descripción del requisito RDC-004

RDC-005	
Prioridad	Alta
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Legibilidad del manual de usuario.

**Tabla 130.** Descripción del requisito RDC-005

RDC-006	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Idioma español.

**Tabla 131.** Descripción del requisito RDC-006

RSG-001	
Prioridad	Alta
Necesidad	Esencial
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Mostrar la contraseña del usuario en la aplicación con máscara de caracteres “*”.

**Tabla 132.** Descripción del requisito RSG-001

RSG-002	
Prioridad	Media
Necesidad	Deseable
Claridad	Alta
Fuente	Cliente
Estabilidad	Estable
Verificabilidad	Alta
Descripción	Aplicar un algoritmo de codificación sobre la contraseña de la cuenta de usuario, cuando se vaya a insertar en el fichero de configuración, de manera que si se edita ese fichero dicha contraseña no sea legible.

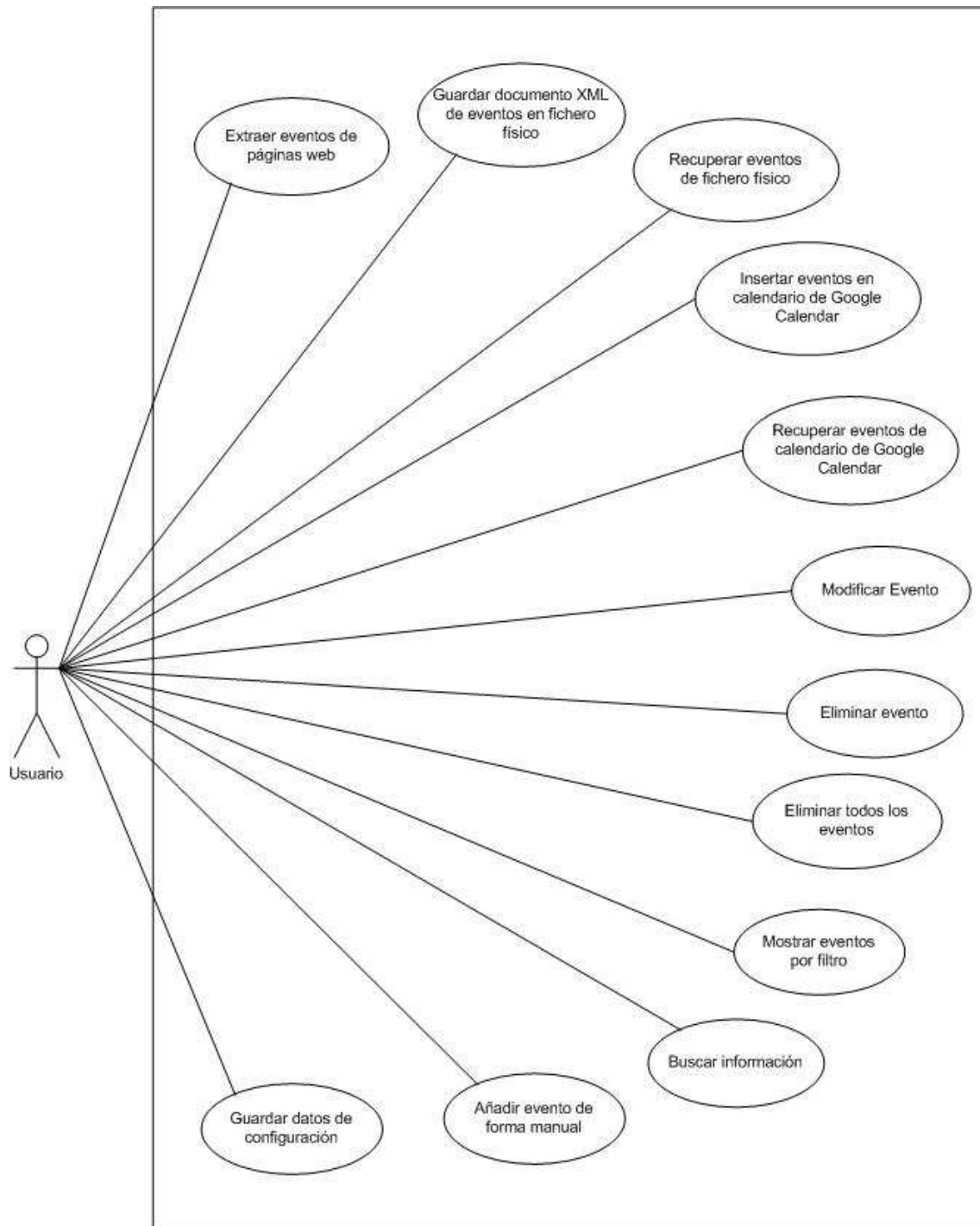
**Tabla 133.** Descripción del requisito RSG-002

## 4.2 DIAGRAMA DE CASOS DE USO

A continuación se muestra una tabla (Tabla 134) con los casos de uso del sistema, el diagrama de casos de uso correspondiente (Figura 18) y, posteriormente, las tablas de descripción de cada caso de uso.

ID	Nombre
CU-001	Extraer eventos de páginas Web
CU-002	Guardar documento XML de eventos en fichero físico
CU-003	Recuperar eventos de fichero físico
CU-004	Insertar eventos en calendario de Google Calendar
CU-005	Recuperar eventos de calendario de Google Calendar
CU-006	Modificar evento
CU-007	Eliminar evento
CU-009	Eliminar todos los eventos
CU-009	Mostrar eventos por filtro
CU-010	Buscar información en páginas Web
CU-011	Añadir nuevo evento de forma manual
CU-012	Actualizar datos de configuración

**Tabla 134.** Tabla de identificación de casos de uso del sistema



**Figura 18.** Diagrama de casos de uso del sistema

## Actores del sistema

Un actor es un rol que un usuario juega con respecto al sistema. Es importante destacar el uso de la palabra rol, pues con esto se especifica que un actor no necesariamente representa a una persona en particular, sino más bien la labor que realiza frente al sistema.

Los actores pueden ser usuarios, otros sistemas, software, hardware o cualquier cosa que interactúe con nuestro sistema.

Los actores fundamentales de este sistema son los denominados “Usuarios” finales de la aplicación, es decir, aquellas figuras que utilizarán y sacarán rendimiento del software desarrollado.

Las tablas que vienen a continuación se corresponden con las descripciones de cada uno de los casos de uso del sistema:

## Descripciones de los casos de uso

CU-001	
Nombre	Extraer eventos de páginas Web
Descripción	Extrae automáticamente los eventos de las páginas Web de conferencias científicas
Actores	Usuario
Precondiciones	Tener conexión a Internet
Flujo	1. Seleccionar la/s página/s Web de eventos 2. Clic sobre el botón de extraer eventos
Postcondiciones	Eventos insertados en documento XML mostrados en la aplicación

*Tabla 135.* Descripción del caso de uso CU-001

CU-002	
Nombre	Guardar documento XML de eventos en fichero físico
Descripción	Almacena el documento XML que contiene los eventos obtenidos de la extracción de información en una ruta física del disco duro
Actores	Usuario
Precondiciones	Existencia del documento XML con los eventos
Flujo	1.Clic sobre el botón de guardar documento XML 2.Seleccionar la ubicación en el disco duro donde se va a guardar el fichero 3.Escribir nombre del fichero 4.Aceptar
Postcondiciones	Fichero creado

**Tabla 136.** Descripción del caso de uso CU-002

CU-003	
Nombre	Recuperar eventos de fichero físico
Descripción	Acceder al fichero del disco duro que contiene los eventos y crear con esto un documento XML
Actores	Usuario
Precondiciones	Existencia del fichero que contiene los eventos
Flujo	1.Clic sobre el botón de recuperar documento XML 2.Seleccionar la ubicación donde se encuentra el fichero 3.Aceptar
Postcondiciones	Documento XML generado con los eventos

**Tabla 137.** Descripción del caso de uso CU-003

CU-004	
Nombre	Insertar eventos en calendario de Google Calendar
Descripción	Insertar los eventos recuperados del proceso de extracción de información en un calendario de Google Calendar.
Actores	Usuario
Precondiciones	Existencia del documento XML que contiene los eventos
Flujo	1.Clic sobre el botón insertar eventos en Google Calendar
Postcondiciones	Eventos insertados en un calendario de Google Calendar

**Tabla 138.** Descripción del caso de uso CU-004

CU-005	
Nombre	Recuperar eventos de calendario de Google Calendar
Descripción	Recuperar los eventos del calendario que contiene los eventos dentro de Google Calendar
Actores	Usuario
Precondiciones	Existencia de eventos en calendario de Google Calendar
Flujo	1.Clic sobre el botón recuperar eventos de calendario de Google Calendar
Postcondiciones	Eventos mostrados en el sistema

**Tabla 139.** Descripción del caso de uso CU-005



CU-006	
Nombre	Modificar evento
Descripción	Modificar los datos de un evento que se encuentre en el sistema y actualizarlo en Google Calendar
Actores	Usuario
Precondiciones	Existencia de eventos en el sistema
Flujo	1.Seleccionar el evento a modificar 2.Seleccionar la opción modificar 3.Modificar datos del evento 4.Aceptar
Postcondiciones	Datos del evento actualizados en el sistema

**Tabla 140.** Descripción del caso de uso CU-006

CU-007	
Nombre	Eliminar evento
Descripción	Eliminar un evento del sistema y del calendario de Google Calendar
Actores	Usuario
Precondiciones	Existencia de eventos en el sistema
Flujo	1.Seleccionar el evento a eliminar 2.Seleccionar la opción eliminar 3.Aceptar aviso de eliminación
Postcondiciones	Evento eliminado del sistema

**Tabla 141.** Descripción del caso de uso CU-007

CU-008	
Nombre	Eliminar todos los eventos
Descripción	Eliminar todos los eventos del sistema y del calendario de Google Calendar
Actores	Usuario
Precondiciones	Existencia de eventos en el sistema
Flujo	1.Clic sobre el botón eliminar todos los eventos 2.Aceptar aviso de eliminación
Postcondiciones	Evento eliminado del sistema

**Tabla 142.** Descripción del caso de uso CU-008

CU-009	
Nombre	Mostrar eventos por filtro
Descripción	Muestra los eventos en el sistema aplicándole un filtro. Las opciones de selección disponibles son: Mostrar todos los eventos, mostrar eventos por día, mostrar eventos por nombre, mostrar eventos por localización, mostrar eventos por rango de fechas
Actores	Usuario
Precondiciones	Existencia de eventos en el sistema
Flujo	1.Seleccionar el filtro
Postcondiciones	Evento mostrados en función del filtro seleccionado

**Tabla 143.** Descripción del caso de uso CU-009

CU-010	
Nombre	Buscar información en páginas Web
Descripción	Buscar información en páginas Web a través del buscador de Google Search incorporado en el sistema
Actores	Usuario
Precondiciones	Ninguna
Flujo	1.Insertar texto a buscar 2.Clic sobre el botón buscar del buscador
Postcondiciones	Páginas Web encontradas

**Tabla 144.** Descripción del caso de uso CU-010

CU-011	
Nombre	Añadir nuevo evento de forma manual
Descripción	Añadir un nuevo eventos en el calendario de Google Calendar insertando los datos de forma manual
Actores	Usuario
Precondiciones	Conexión a Internet
Flujo	1.Insertar los datos del nuevo evento 2.Clic en el botón Aceptar
Postcondiciones	Páginas Web encontradas

**Tabla 145.** Descripción del caso de uso CU-011

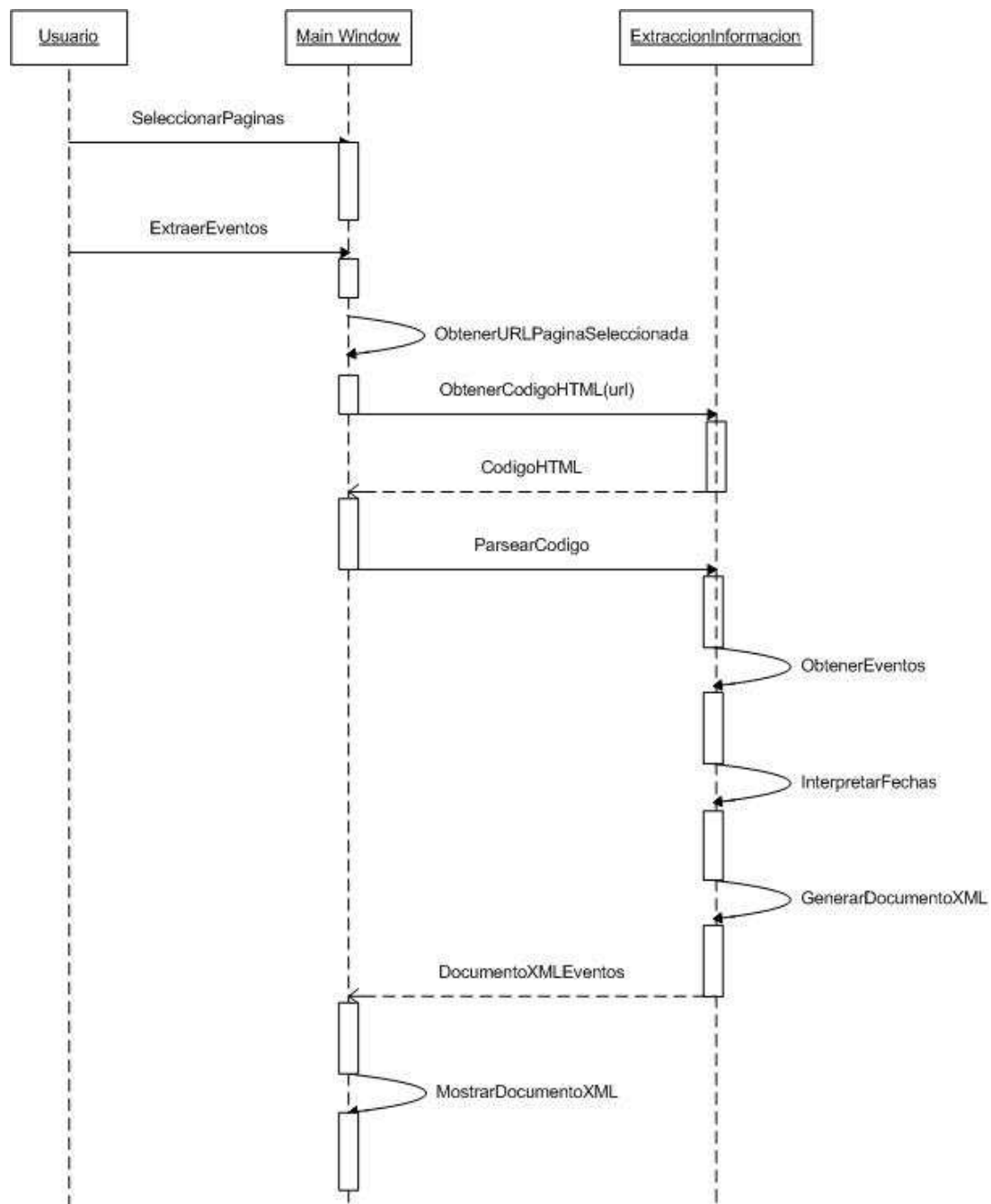
CU-012	
Nombre	Actualizar datos de configuración
Descripción	Guarda los datos en el fichero de configuración del sistema. Este caso de uso agrupan las distintas partes: Datos de acceso a Google para un usuario, fechas de recuperación de eventos, opciones de proxy, nombre del calendario de eventos de Google
Actores	Usuario
Precondiciones	Datos existentes en el fichero de configuración recuperados en el sistema
Flujo	1.Insertar los datos a actualizar 2.Guardar
Postcondiciones	Nuevos datos guardados en el fichero

**Tabla 146.** Descripción del caso de uso CU-012

### 4.3 DIAGRAMAS DE SECUENCIA

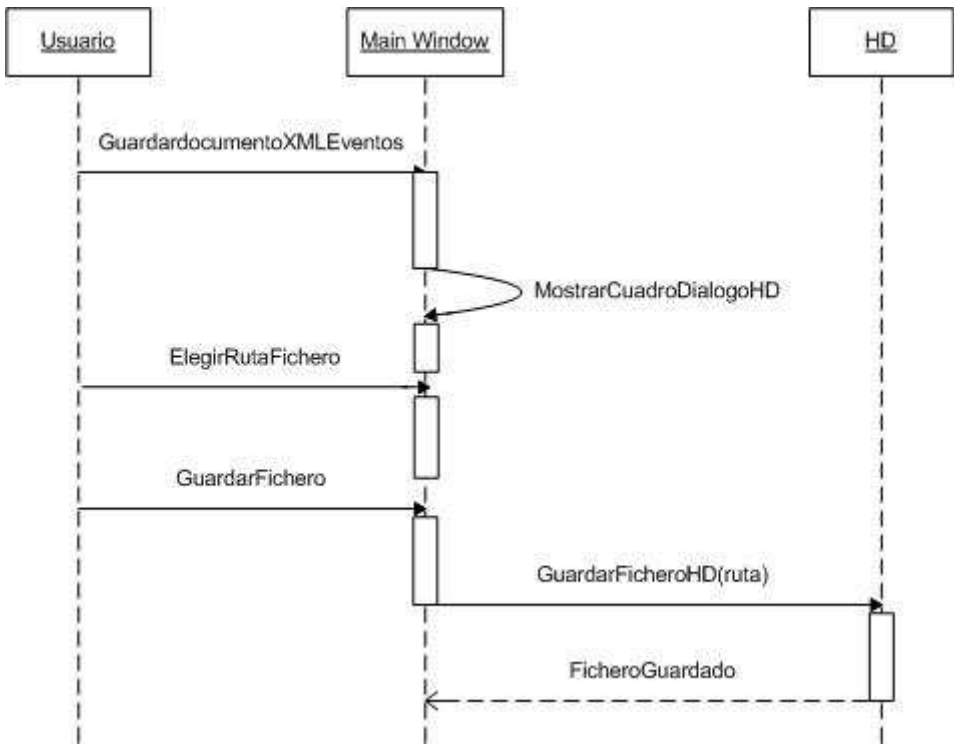
A continuación se muestran los diagramas de secuencia aplicados a cada caso de uso desarrollado en el punto anterior.

#### CU-001 Extraer eventos de páginas Web (Figura 19):



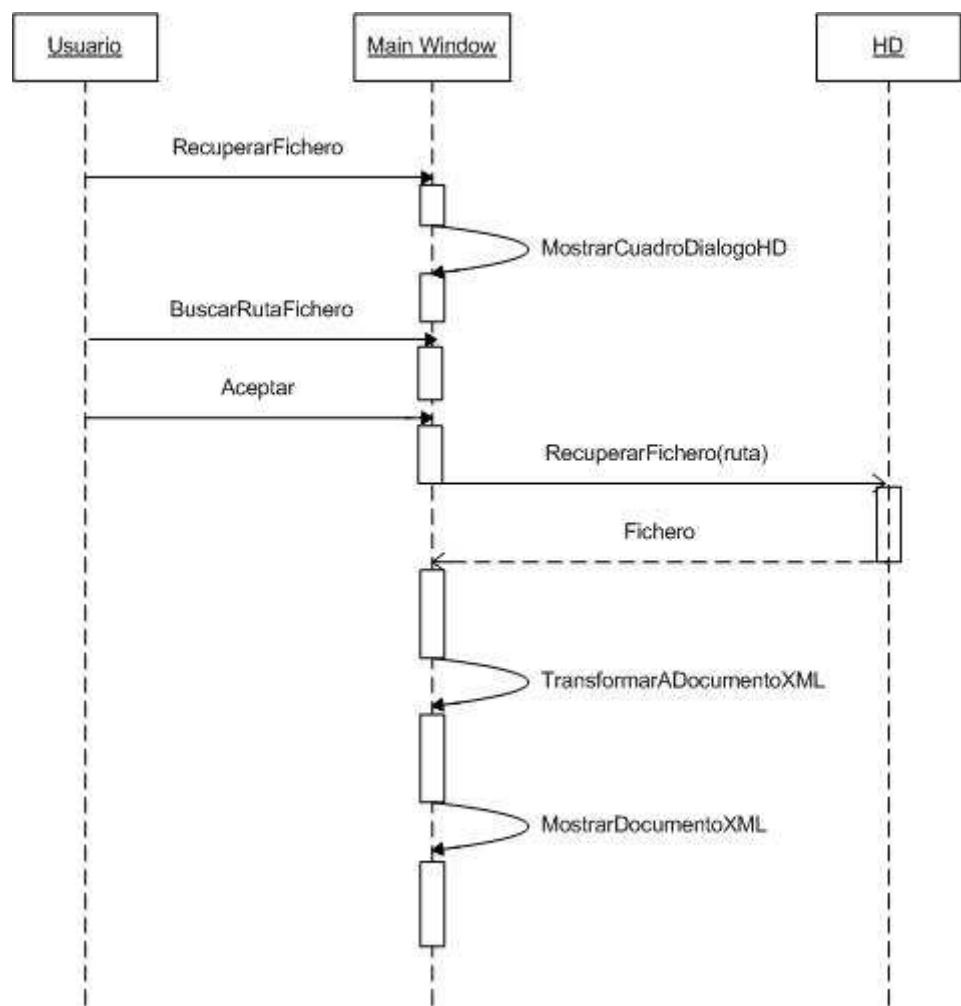
**Figura 19.** Diagrama de secuencia que describe el caso de uso CU-001

**CU-002 Guardar documento XML de eventos en fichero físico (Figura 20):**



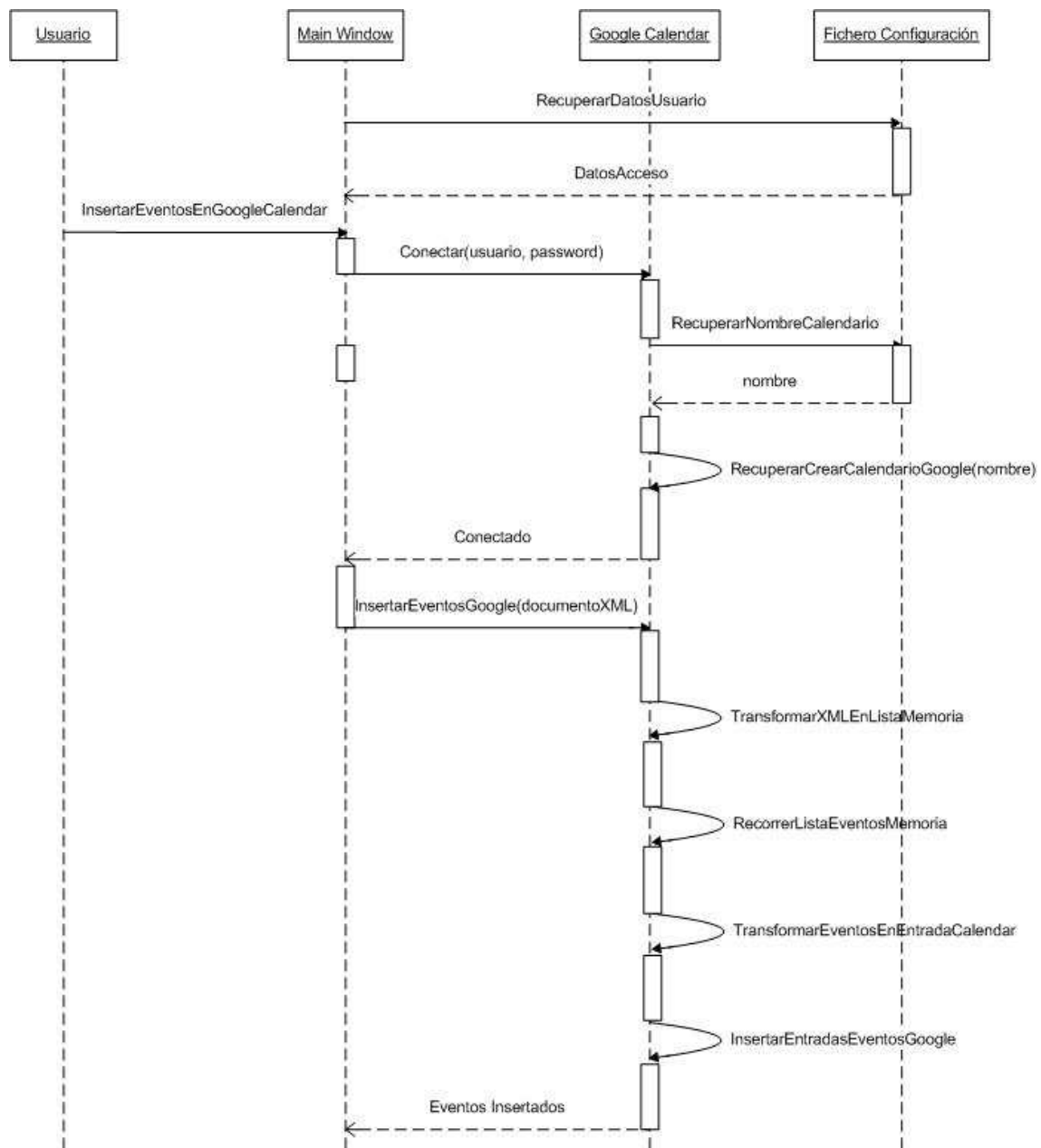
*Figura 20.* Diagrama de secuencia que describe el caso de uso CU-002

**CU-003 Recuperar eventos de fichero físico (Figura 21):**



**Figura 21.** Diagrama de secuencia que describe el caso de uso CU-003

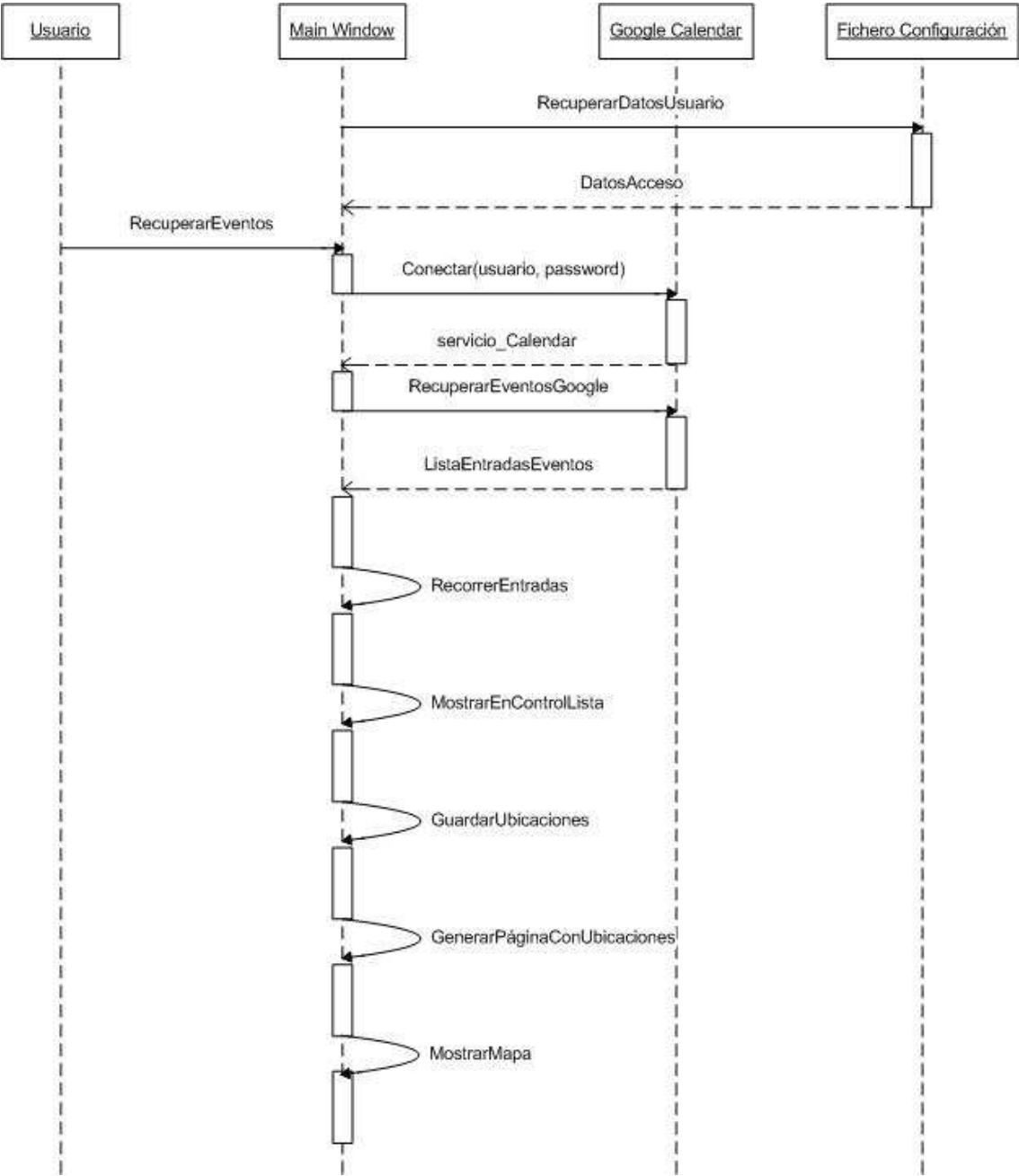
**CU-004 Insertar eventos en calendario de Google Calendar (Figura 22):**



**Figura 22.** Diagrama de secuencia que describe el caso de uso CU-004

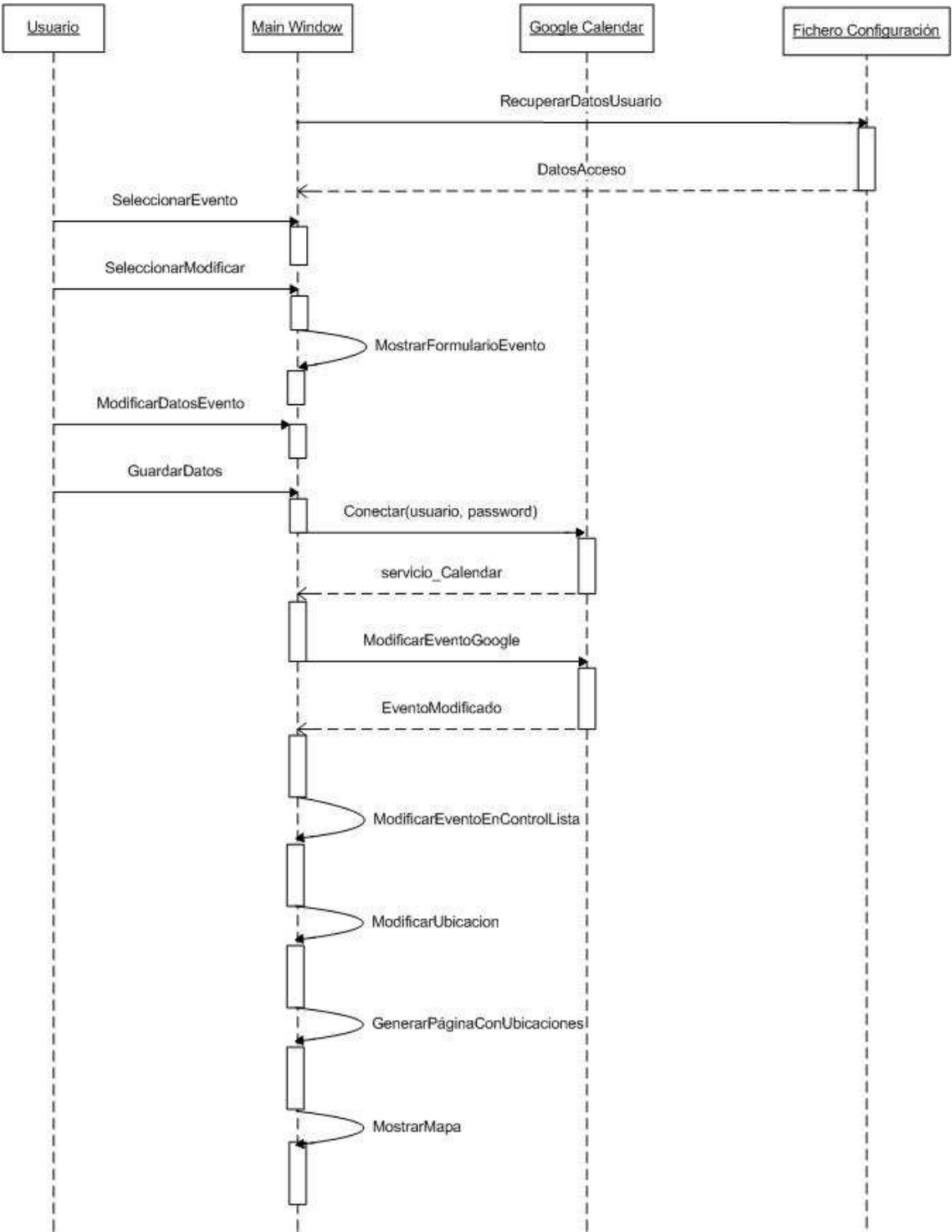


**CU-005 Recuperar eventos de calendario de Google Calendar (Figura 23):**



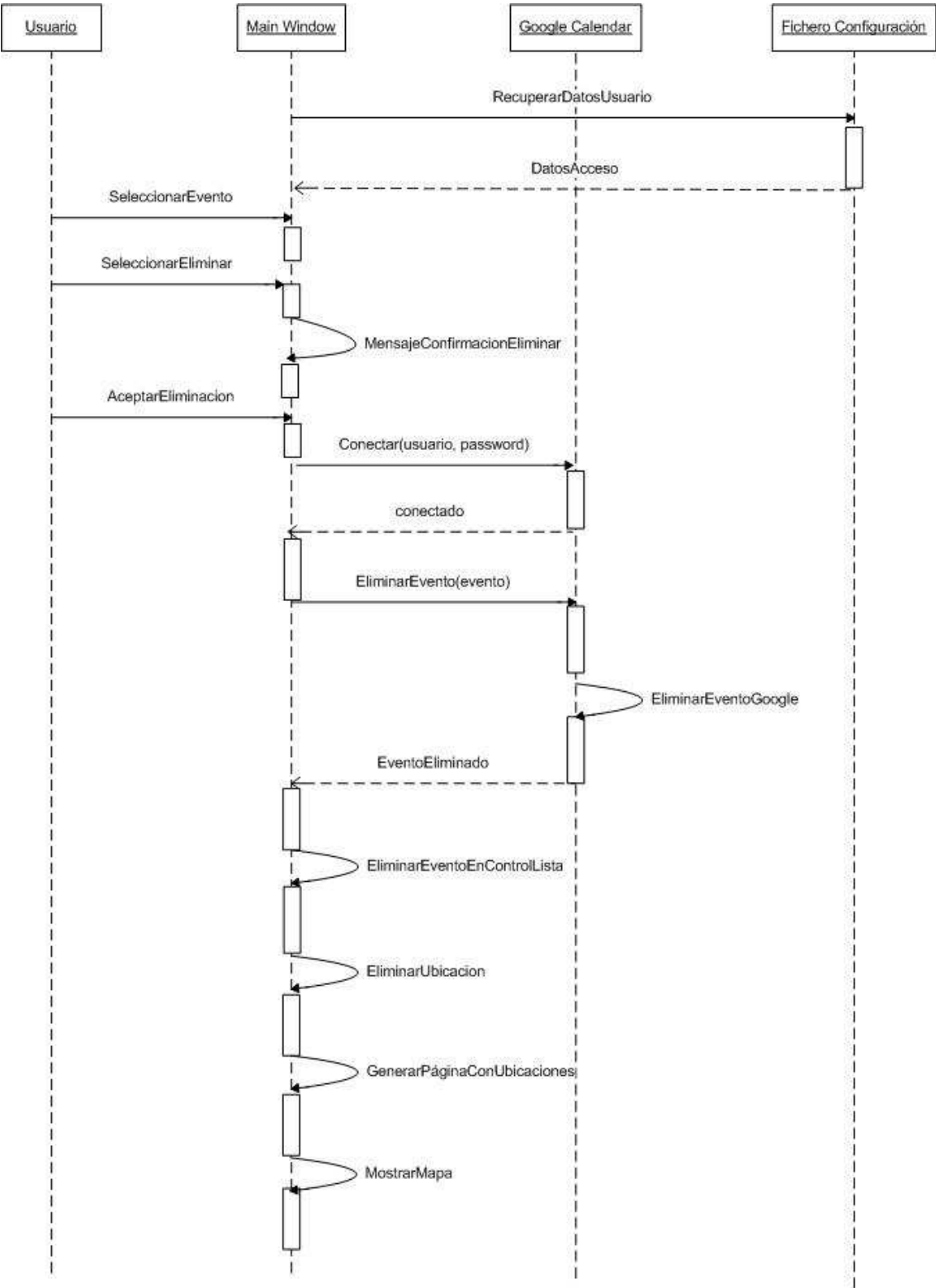
**Figura 23.** Diagrama de secuencia que describe el caso de uso CU-005

**CU-006 Modificar evento (Figura 24):**



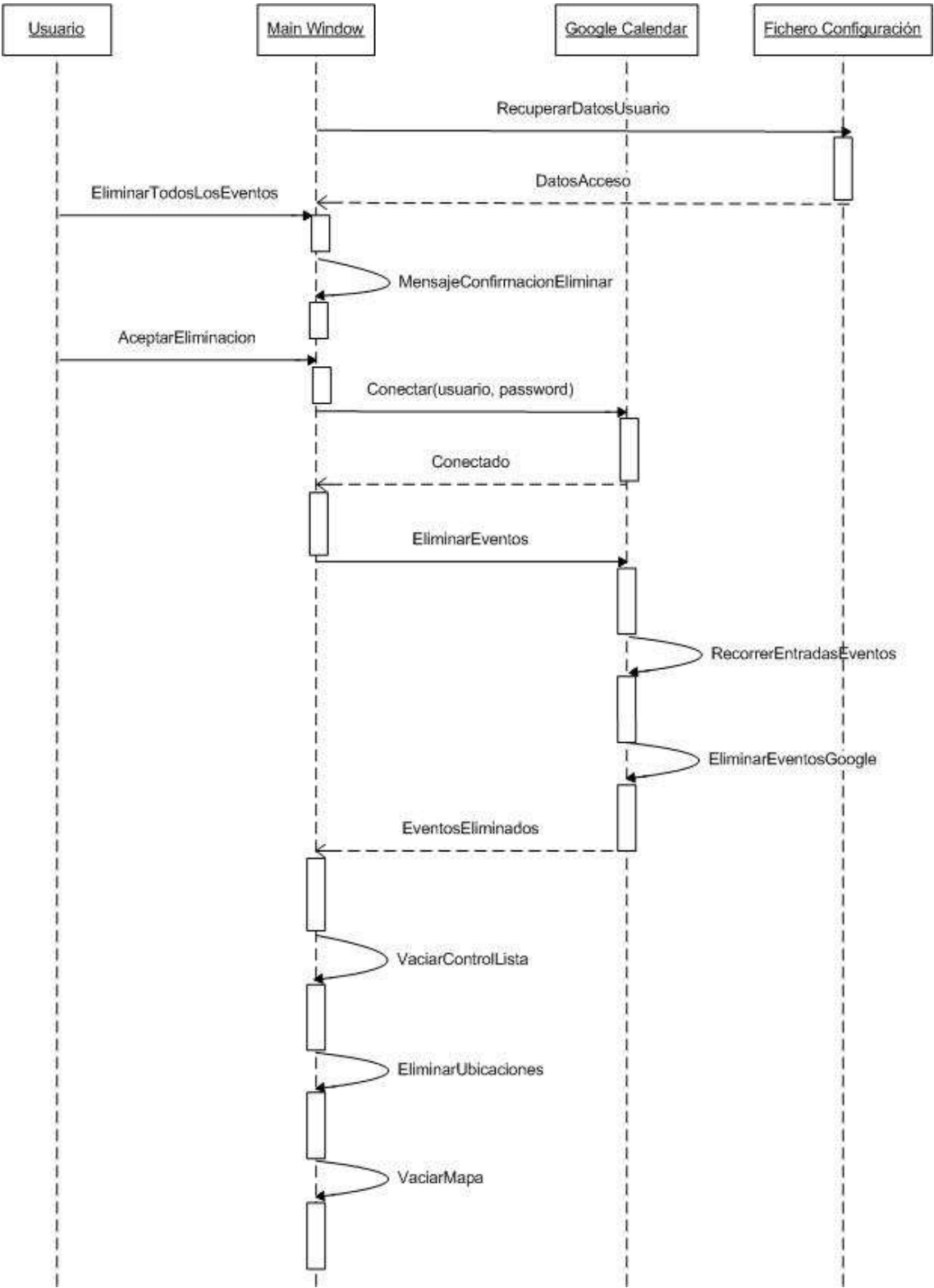
*Figura 24.* Diagrama de secuencia que describe el caso de uso CU-006

**CU-007 Eliminar evento (Figura 25):**



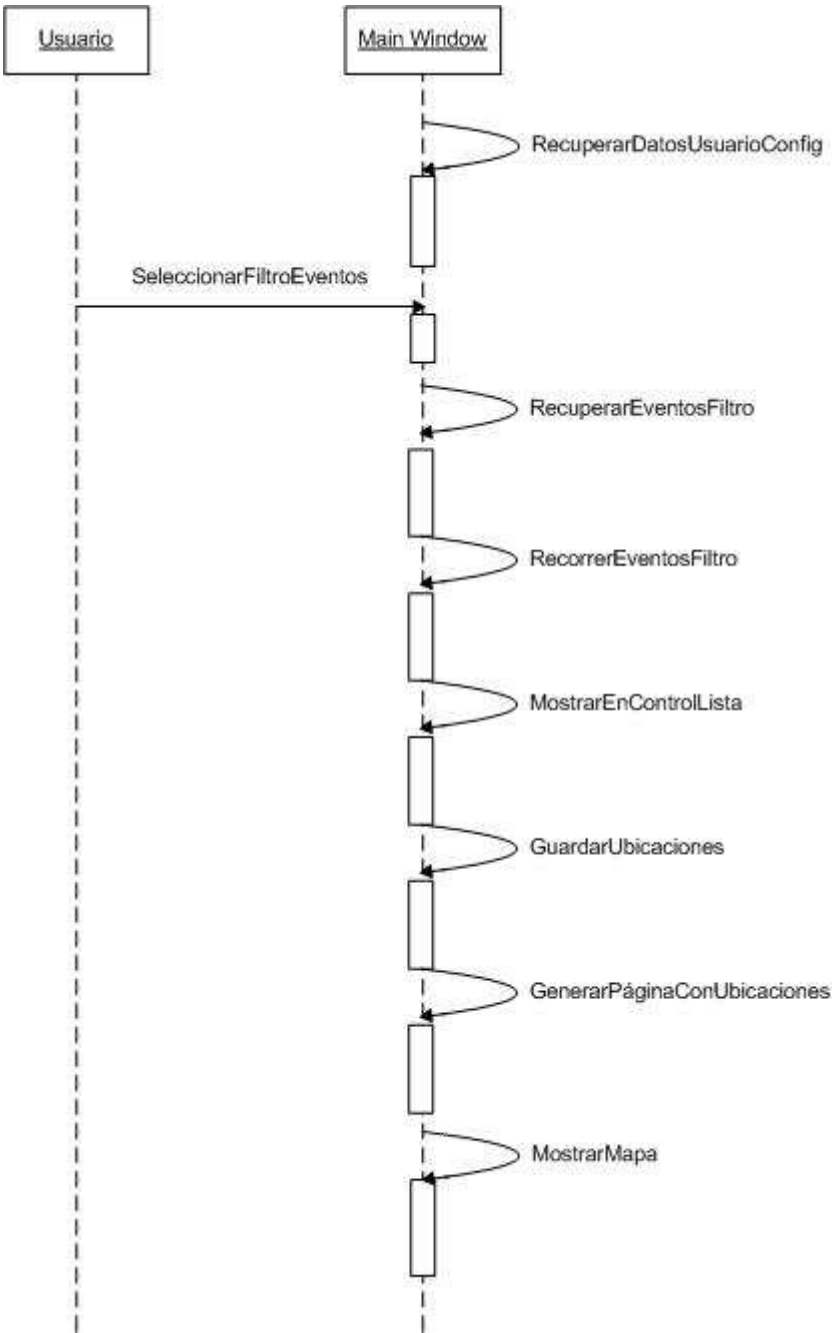
*Figura 25.* Diagrama de secuencia que describe el caso de uso CU-007

**CU-008 Eliminar todos los eventos (Figura 26):**



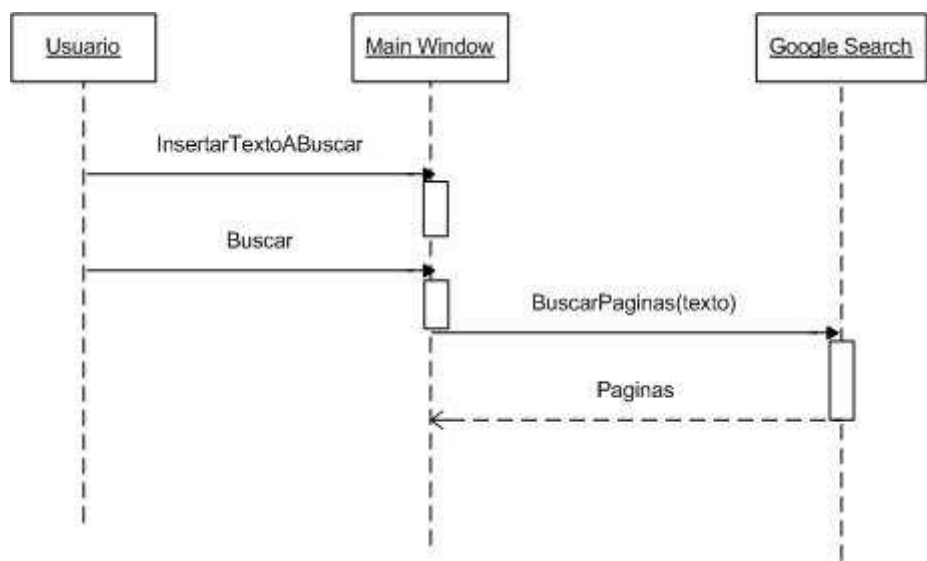
*Figura 26.* Diagrama de secuencia que describe el caso de uso CU-008

**CU-009 Mostrar eventos por filtro (Figura 27):**



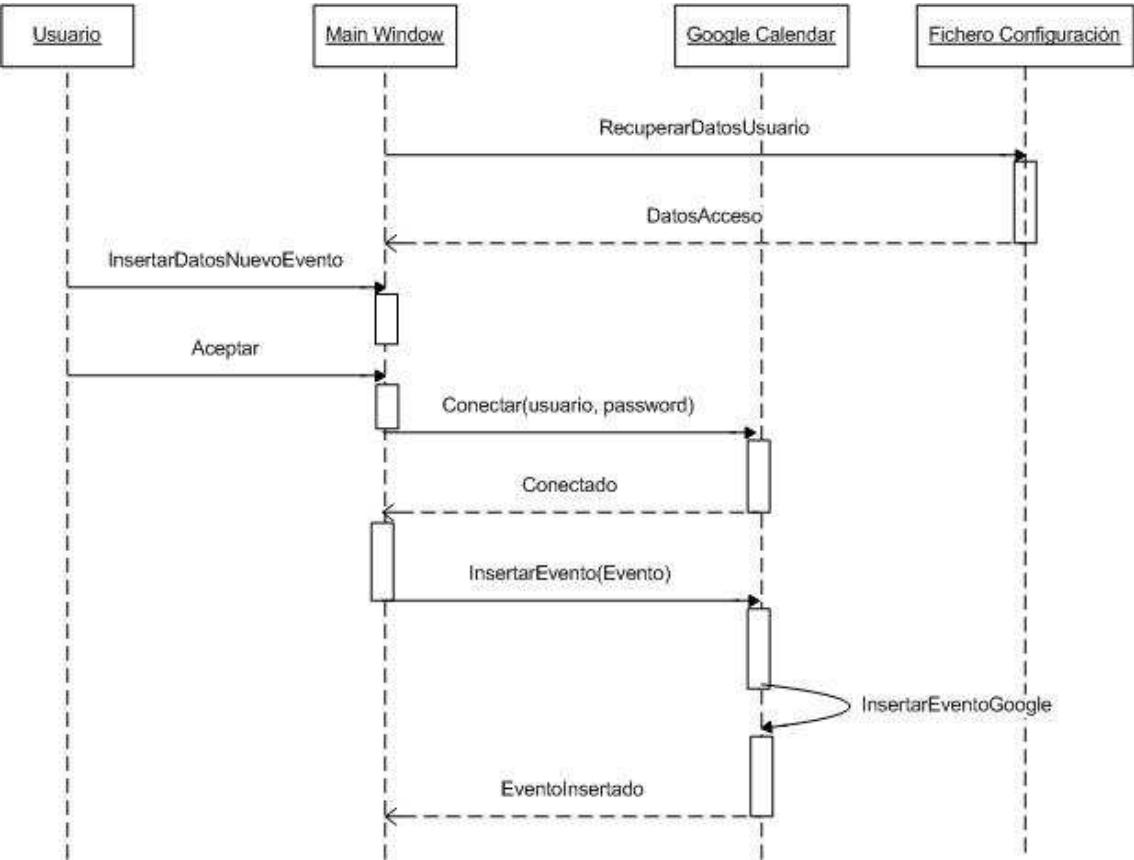
**Figura 27.** Diagrama de secuencia que describe el caso de uso CU-009

**CU-010 Buscar información en páginas Web (Figura 28):**



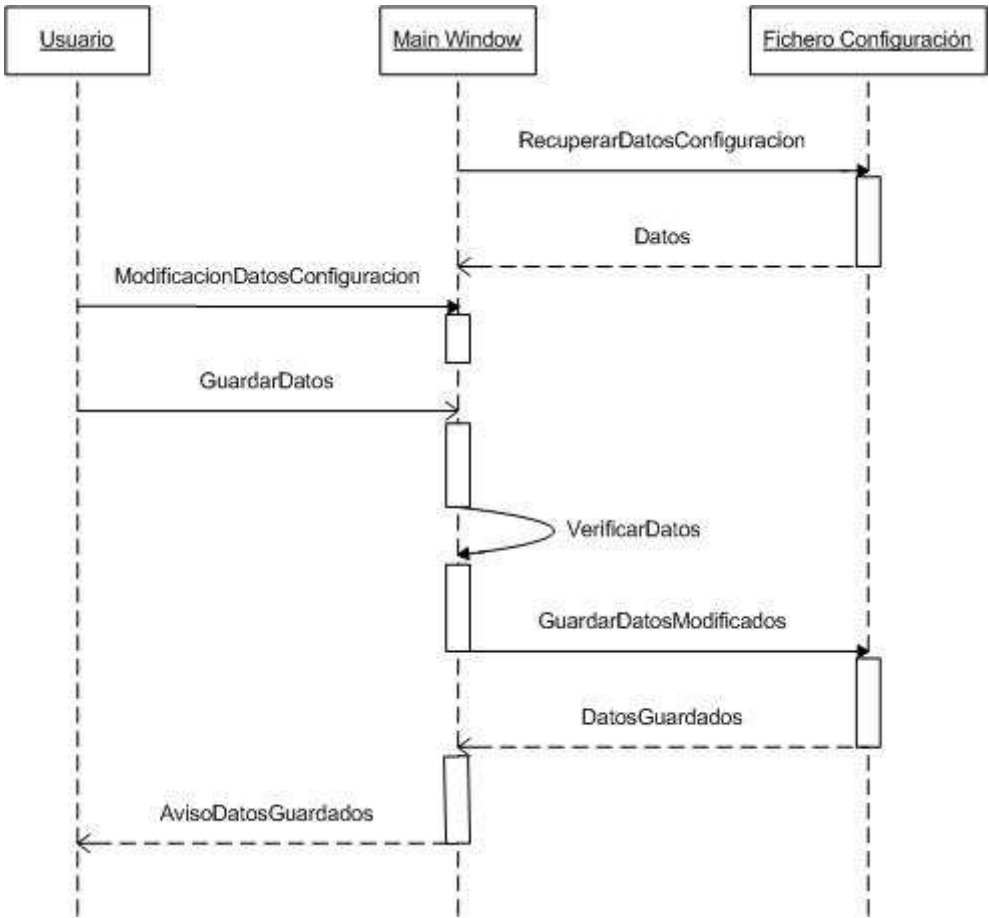
*Figura 28.* Diagrama de secuencia que describe el caso de uso CU-010

**CU-011 Añadir nuevo evento de forma manual (Figura 29):**



*Figura 29.* Diagrama de secuencia que describe el caso de uso CU-011

**CU-012 Actualizar datos de configuración (Figura 30):**



**Figura 30.** Diagrama de secuencia que describe el caso de uso CU-012



#### 4.4 DIAGRAMA DE CLASES

A continuación se muestra el diagrama de clases del sistema con sus propiedades y métodos más importantes (Figura 31).

Los siguientes puntos de este apartado contienen la descripción de las clases, propiedades y métodos públicos que se encuentran en ellas.

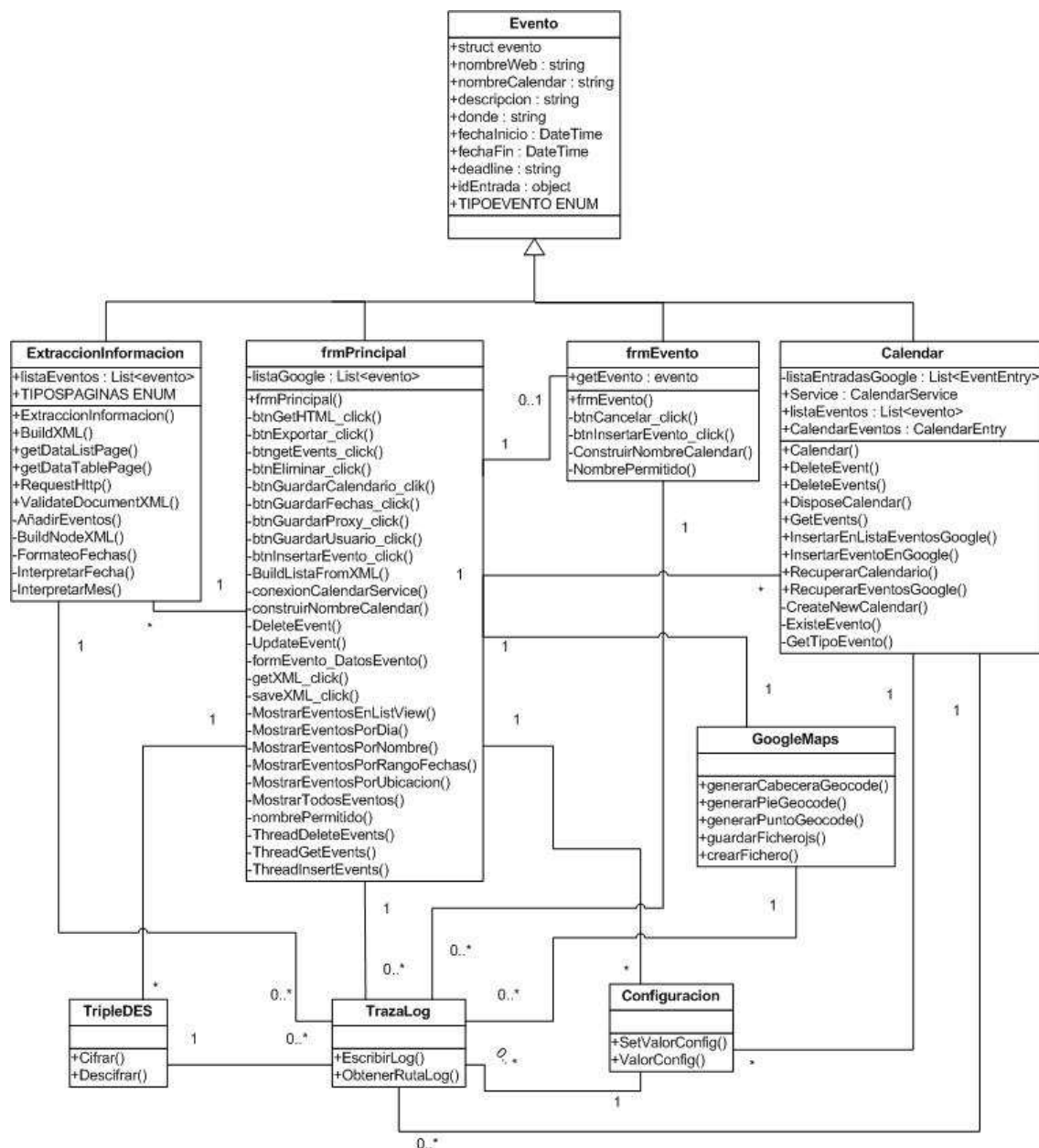


Figura 31. Diagrama de clases del sistema

#### 4.4.1 DESCRIPCIÓN DE LAS CLASES DEL SISTEMA

La siguiente tabla (Tabla 147) muestra la descripción de las clases definidas en el sistema:

Clase	Descripción
Definiciones	Clase que contiene la estructura y datos de definición para almacenar un evento en el sistema.
ExtraccionInformacion	Clase que contiene toda la funcionalidad del proceso de extracción de información de forma automática.
Calendar	Clase que contiene la conexión y todos los métodos para acceder e interactuar con Google Calendar.
Configuracion	Clase de acceso a ficheros de configuración de una aplicación.
frmEvento	Clase formulario de edición de un evento.
frmPrincipal	Clase que contiene el interfaz principal de la aplicación.
gMaps	Clase que construye todos los puntos de ubicación de eventos en JavaScript para incrustarlo en una página HTML.
Program	Clase donde se encuentra el método main(), que arranca la aplicación.
TrazaLog	Clase que escribe los mensajes de error en el fichero de LOG.
TripleDES	Clase de cifrado con codificación Triple DES.

**Tabla 147.** Descripción de las clases del sistema

## 4.4.2 CLASE DEFINICIONES

### 4.4.2.1 Propiedades de la clase Definiciones

La siguiente tabla (Tabla 148) muestra la descripción de las propiedades públicas definidas dentro de la clase Definiciones:

Propiedad	Descripción
Evento	Estructura con la información recuperada de un evento.
TIPOSEVENTOS	Enumerado con los tipos de eventos que pueden existir.

*Tabla 148.* Descripción de las propiedades de la clase Definiciones

La siguiente tabla (Tabla 149) muestra las propiedades contenidas en la estructura *Evento*:

Propiedad	Descripción
DeadLine	Propiedad que almacena la fecha de deadline del evento.
Descripcion	Propiedad que almacena la descripción del evento.
Donde	Propiedad que almacena el lugar donde tendrá lugar el evento.
FechaFin	Propiedad que almacena la fecha de finalización del evento.
FechaInicio	Propiedad que almacena la fecha inicio del evento.
IdEntrada	Propiedad que almacena el identificador de entrada del evento en Google Calendar.
NombreCalendar	Propiedad que almacena el nombre del evento generado para ser insertado en el calendario de Google Calendar.

NombreWeb	Propiedad que almacena el nombre del evento cuando se recupera de la página Web.
TipoEvento	Propiedad que almacena el tipo de evento.

*Tabla 149.* Descripción de las propiedades que contiene la estructura Evento

La siguiente tabla (Tabla 150) muestra los valores de TIPOEVENTOS:

Tipo Evento	Descripción
Mainconference	Tipo de evento Main Conference
Deadline	Tipo de evento DeadLine
Submission	Tipo de evento DeadLine - Full Submission
Workshops	Tipo de evento DeadLine – WorkShops
Abstracts	Tipo de evento DeadLine – Abstracts

*Tabla 150.* Descripción de los valores de TIPOSEVENTOS

#### 4.4.2.2 Métodos de la clase Definiciones

La siguiente tabla (Tabla 151) muestra la descripción de los métodos definidos dentro de la clase Definiciones:

Método	Descripción
Definiciones Constructor	Inicializa una nueva instancia de la clase Definiciones.

*Tabla 151.* Descripción de los métodos de la clase Definiciones

### 4.4.3 CLASE EXTRACCIONINFORMACION

#### 4.4.3.1 Propiedades públicas de la clase ExtraccionInformacion

La siguiente tabla (Tabla 152) muestra la descripción de las propiedades públicas definidas dentro de la clase ExtraccionInformacion:

Propiedad	Descripción
DocEvents	Propiedad que devuelve el documento XML de eventos.
ListaEventos	Propiedad que devuelve la lista de eventos.
TIOSPPAGINAS	Enumerado con los tipos de páginas que se pueden leer e interpretar.

**Tabla 152.** Descripción de las propiedades de la clase ExtraccionInformacion

La siguiente tabla (Tabla 153) muestra los tipos de páginas que se encuentran dentro de TIOSPPAGINAS:

Tipo Página	Descripción
WikiCFP	Página tipo WikiCFP
ACLWiki	Página tipo ACLWiki

**Tabla 153.** Descripción de los valores de TIOSPPAGINAS

#### 4.4.3.2 Métodos de la clase ExtraccionInformacion

La siguiente tabla (Tabla 154) muestra la descripción de los métodos definidos dentro de la clase ExtraccionInformacion:

Método	Descripción
ExtraccionInformacion Constructor	Constructor de la clase. Inicializa la lista de eventos y el documento XML de eventos.
BuildXML	Construye el documento XML a través de los objetos evento insertados en la lista de eventos.
getAllNodes	Ejemplo de uso del HTMLParser.

getDataListPage	Se recorre las páginas de tipo ACLWiki, interpreta los eventos y guarda los datos en un objeto <i>List</i> (lista de eventos).
getDataTablePage	Se recorre las páginas de tipo WikiCFP, interpreta los eventos y guarda los datos en un objeto <i>List</i> (lista de eventos).
RequestHttp	Método que obtiene el código fuente de una Web a través de una petición Request de Http. Si los parámetros referentes al proxy llegan vacíos intenta conectar sin proxy.
ValidateDocumentXML	Otro método de validación de documentos XML.

**Tabla 154.** Descripción de los métodos de la clase ExtraccionInformacion

#### 4.4.4 CLASE CALENDAR

##### 4.4.4.1 Propiedades públicas de la clase Calendar

La siguiente tabla (Tabla 155) muestra la descripción de las propiedades públicas definidas dentro de la clase Calendar:

Propiedad	Descripción
calendarEventos	Devuelve la entrada del calendario de eventos creado en Google Calendar.
listaEventosGoogle	Devuelve la lista con los eventos recuperados del calendario de Google Calendar.
service	Devuelve la variable con el servicio de Google.

**Tabla 155.** Descripción de las propiedades públicas de la clase Calendar

#### 4.4.4.2 Métodos de la clase Calendar

La siguiente tabla (Tabla 156) muestra la descripción de los métodos definidos dentro de la clase Calendar:

Método	Descripción
Calendar Constructor	Constructor de la clase Calendar. Inicializa las variables con valores del fichero de configuración. Recupera el identificador del calendario o lo crea si no existe.
AddSubscription	Añade una suscripción al calendario.
CreateEventQuickAdd	Crea un evento QuickAdd.
DeleteCalendar	Elimina todos los calendarios de Google Calendar.
DeleteEvent	Elimina un evento concreto.
DeleteEvents	Eliminan todos los eventos en Google Calendar. Y en las listas de eventos asociadas.
DeleteSubscriptions	Elimina una suscripción del calendario.
DisposeCalendar	Vacía todas las variables.
getAllCalendars	Recupera todos los calendarios creado en Google a través del servicio de Google Calendar.
InsertarEnListaEventosGoogle	Inserta la entrada del evento en la lista de eventos de Google.
InsertarEventoEnGoogle	Inserta la entrada de un evento en el calendario de Google Calendar.
InsertarEventosEnGoogle	Inserta en un calendario de Google los eventos incluidos en la lista de eventos ( <i>List</i> ).
InsertEvents	Inserta los eventos que se encuentran incluidos en el documento de eventos.
RecuperarCalendario	Recupera el calendario que coincida con el nombre de calendario, que se le pasa como parámetro.

<b>Método</b>	<b>Descripción</b>
RecuperarEventosGoogle	Recupera los eventos que se encuentran en el calendario de la aplicación en Google Calendar.
RecuperarEventosLocalizacion	Recuperan los eventos para una localización concreta de la lista de eventos de Google.
RecuperarEventosNombre	Recupera los eventos para un rango de fechas de la lista de eventos de Google.
RecuperarEventosRangoFechas	Recupera los eventos para un rango de fechas de la lista de eventos de Google.
RecuperarLocalizaciones	Recupera las localizaciones de todos los eventos.
UpdateCalendar	Actualiza las propiedades especificadas en el calendario de Google Calendar.
UpdateEvent	Actualiza un evento.
UpdateSuscriptions	Actualiza el color de las entradas del calendario.
Finalize	Destructor de la clase Calendar.

*Tabla 156.* Descripción de los métodos de la clase Calendar



#### 4.4.5 CLASE CONFIGURACION

##### 4.4.5.1 Métodos de la clase Configuracion

La siguiente tabla (Tabla 157) muestra la descripción de los métodos definidos dentro de la clase Configuracion:

Método	Descripción
Configuracion Constructor	Inicializa una nueva instancia de la clase Configuracion.
SetValorConfig	Modifica el valor de una clave de la sección AppSettings del fichero de configuración.
ValorConfig	Obtiene un valor de la clave de una sección fichero de configuración.

*Tabla 157.* Descripción de los métodos de la clase Configuracion

#### 4.4.6 CLASE FRMEVENTO

##### 4.4.6.1 Propiedades públicas de la clase FrmEvento

La siguiente tabla (Tabla 158) muestra la descripción de las propiedades públicas definidas dentro de la clase FrmEvento:

Propiedad	Descripción
getEvento	Propiedad que devuelve el evento cargado en el formulario.

*Tabla 158.* Descripción de las propiedades de la clase FrmEvento

##### 4.4.6.2 Métodos de la clase FrmEvento

La siguiente tabla (Tabla 159) muestra la descripción de los métodos definidos dentro de la clase FrmEvento:

Método	Descripción
frmEvento	Inicializa una nueva instancia de la clase frmEvento.
Dispose	Vacía los recursos que se han usado

*Tabla 159.* Descripción de los métodos de la clase FrmEvento

#### 4.4.7 CLASE FRMPRINCIPAL

##### 4.4.7.1 Métodos de la clase frmPrincipal

La siguiente tabla (Tabla 160) muestra la descripción de los métodos definidos dentro de la clase FrmPrincipal:

Método	Descripción
frmPrincipal Constructor	Constructor del formulario. Inicializa los componentes.
Dispose	Vacía los recursos que se han usado.

*Tabla 160.* Descripción de los métodos de la clase FrmPrincipal

#### 4.4.8 CLASE GMAPS

##### 4.4.8.1 Métodos de la clase gMaps

La siguiente tabla (Tabla 161) muestra la descripción de los métodos definidos dentro de la clase gMaps:

<b>Método</b>	<b>Descripción</b>
gMaps Constructor	Inicializa una nueva instancia de la clase gMaps.
crearFichero	Crea el fichero de JavaScript con la construcción del mapa.
generarCabeceraGeocode	Inicializa variables, y crea el contenedor del mapa.
generarPieGeocode	Crea el pie del punto Geocode (cierres de JavaScript).
generarPuntoGeoCode	Calcula el punto GeoCode para la ubicación, creando el marcador para ese punto, con el evento clic para mostrar el texto asociado (eventos + ubicación).
generarPuntoGeocodeL	Genera un punto de GeoCode al que se le puede indicar el icono que se le va a poner.
guardarFichero	Inserta el script con el mapa, en el fichero ".js".

*Tabla 161.* Descripción de los métodos de la clase gMaps

#### 4.4.9 CLASE TRAZALOG

##### 4.4.9.1 Métodos de la clase TrazaLog

La siguiente tabla (Tabla 162) muestra la descripción de los métodos definidos dentro de la clase TrazaLog:

<b>Método</b>	<b>Descripción</b>
TrazaLog Constructor	Inicializa una nueva instancia de la clase TrazaLog.
EscribirLog	Escribe una entrada en el fichero de log de la aplicación.

*Tabla 162.* Descripción de los métodos de la clase TrazaLog

#### 4.4.10 CLASE TRIPLEDES

#### 4.4.10.1 Métodos de la clase TripleDES

La siguiente tabla (Tabla 163) muestra la descripción de los métodos definidos dentro de la clase TripleDES:

Método	Descripción
TripleDES Constructor	Inicializa una nueva instancia de la clase TripleDES.
Cifrar	Cifrado Triple DES.
Descifrar	Descifrado Triple DES.

*Tabla 163.* Descripción de los métodos de la clase TripleDES

## 5 IMPLEMENTACIÓN DEL SISTEMA

En esta fase se describe como se ha llevado a cabo la construcción del sistema para cumplir los requisitos especificados, y siguiendo los esquemas de la etapa de análisis y diseño.

Para organizar las cuatro partes especificadas en la fase de análisis, se ha utilizado un control de .NET con cuatro pestañas:

1. Pestaña de Extracción de Información Web.
2. Pestaña de Gestión de Eventos de Google.
3. Pestaña de Búsqueda e Inserción Manual de Eventos.
4. Pestaña de Configuración.

### 5.1 PESTAÑA DE EXTRACCIÓN DE INFORMACIÓN

Esta primera pestaña que agrupa la funcionalidad asociada a la fase de extracción de información automática de las páginas Web de eventos de conferencias científicas.

Se han elegido dos tipos de páginas de conferencias científicas como muestra para realizar la extracción automática de eventos. El usuario puede elegir si obtener los eventos de una, otra o las dos a la vez.

El sistema recupera el código fuente de estas páginas. Para ello utilizan dos objetos: *HttpWebRequest* y *HttpWebResponse*, al que se le configuran unos parámetros y permite recuperar el código fuente de una URL.

El sistema está desarrollado de manera modular de forma que sería sencillo ampliar la extracción de información desde nuevas páginas de conferencias, añadiendo los nuevos métodos de parseo en el módulo de extracción de información y haciendo pequeñas modificaciones en la interfaz de la aplicación.

El código obtenido es parseado para obtener los eventos de las conferencias.

Los datos que se recuperan para cada evento son:

- Nombre del evento
- Descripción
- Lugar
- Fecha de Inicio
- Fecha de Fin
- DeadLine

El parseo se realiza utilizando expresiones regulares. No obstante, se han estudiado la utilización de varias posibilidades:

- Utilizar la librería HTMLParser.
- Realizar el parseo utilizando una gramática.
- Utilizar expresiones regulares.

La **librería HTMLParser**: es una librería de código libre que transforma el código HTML en código XML, más fácil de gestionar. En las pruebas que se hicieron, se encontraron varios problemas:

- Si el HTML no está bien formado, puede omitir etiquetas.
- La transformación del HTML en XML lleva tiempo, ralentiza el sistema.
- Existen problemas para obtener datos de tablas anidadas (en uno de los tipo de páginas de eventos, los eventos se encuentran en una tabla anidada dentro de otras).
- El código de las páginas es grande, hay problemas para acceder a las etiquetas anidadas que contienen cada parte de código que se necesita.
- Para obtener los datos hay que hacer otro tipo de parseo de texto.

Utilizar una **gramática para realizar el parseo**: el parseo manual es el método más difícil, y más problemático. En estas páginas el código fuente es extenso, y sería lento y complicado hacer el parseo de toda la información para interpretar la parte que más interesa.

**Expresiones Regulares:** Hay que conocer bien la página que se va a parsear. Se accede a través de un patrón que acota directamente el texto que se quiere recuperar, de forma rápida. Lo complicado de este método es crear el patrón que devuelve la información de forma precisa, sobre todo con las tablas anidadas.

De los tres métodos el que conlleva un menor grado de complejidad y el que requiere menos tiempo de acceso a los datos es la utilización de expresiones regulares, es por ello que se ha optado por éste para realizar el parseo.

Los eventos que se recuperan del proceso de extracción se almacenan en un documento XML. Estos eventos muestran la información tal y como aparece en la página Web, posteriormente deberán ser tratados para insertarlos en Google Calendar.

Para asegurar que el documento XML está bien formado se valida contra el siguiente esquema de validación XSD (Tabla 164):

```

<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
elementFormDefault="qualified">
  <xs:element name="Eventos">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="Evento" minOccurs="0"
maxOccurs="unbounded">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="NombreWeb" type="xs:string" />
              <xs:element name="NombreCalendar" type="xs:string" />
              <xs:element name="Descripcion" type="xs:string" />
              <xs:element name="FechaInicio" type="xs:string" />
              <xs:element name="FechaFin" type="xs:string" />
              <xs:element name="Donde" type="xs:string" />
              <xs:element name="DeadLine" type="xs:string" />
              <xs:element name="TipoEvento" type="xs:string" />
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>

```

**Tabla 164.** Esquema de validación (XSD) aplicado al documento XML de eventos

Este documento XML validado, se le mostrará al usuario en una caja de texto, una vez acabado el parseo. También se mostrará un resumen del resultado de la extracción.

En esta pestaña también se encuentra el botón para insertar los eventos obtenidos en un calendario en Google Calendar.



Con cada operación que realiza el sistema, en todas las pestañas, para interactuar con Google Calendar se realizan las siguientes operaciones:

- Comprobar si se ha accedido previamente a Google Calendar, si se ha accedido se recupera el servicio. Si no se ha accedido antes se conecta a través del servicio que facilita el API. Este servicio requiere los datos de acceso del usuario a su cuenta del gesto de correo de Google (nombre del usuario y contraseña). Estos datos se establecen en la pestaña de configuración.
- Una vez que se ha conectado se comprueba si se ha creado el calendario de evento. Este dato también se establece en la pestaña de configuración. Si el calendario ya existe en Google Calendar se recupera el identificador, si no existiera se procede a la creación recuperando también el identificador del calendario.
- Una vez recuperado el identificador del calendario se procede a insertar los eventos recuperados en la fase de extracción. Estos eventos tienen que ser tratados, antes de insertarlos en el calendario de Google Calendar. Esta operación requiere un nuevo parseo de fechas. En una de las páginas las distintas fechas del evento están agrupadas. Hay que separarlas, creando un evento por cada fecha y modificando para distinguirlos, el nombre del evento concatenando su tipo.

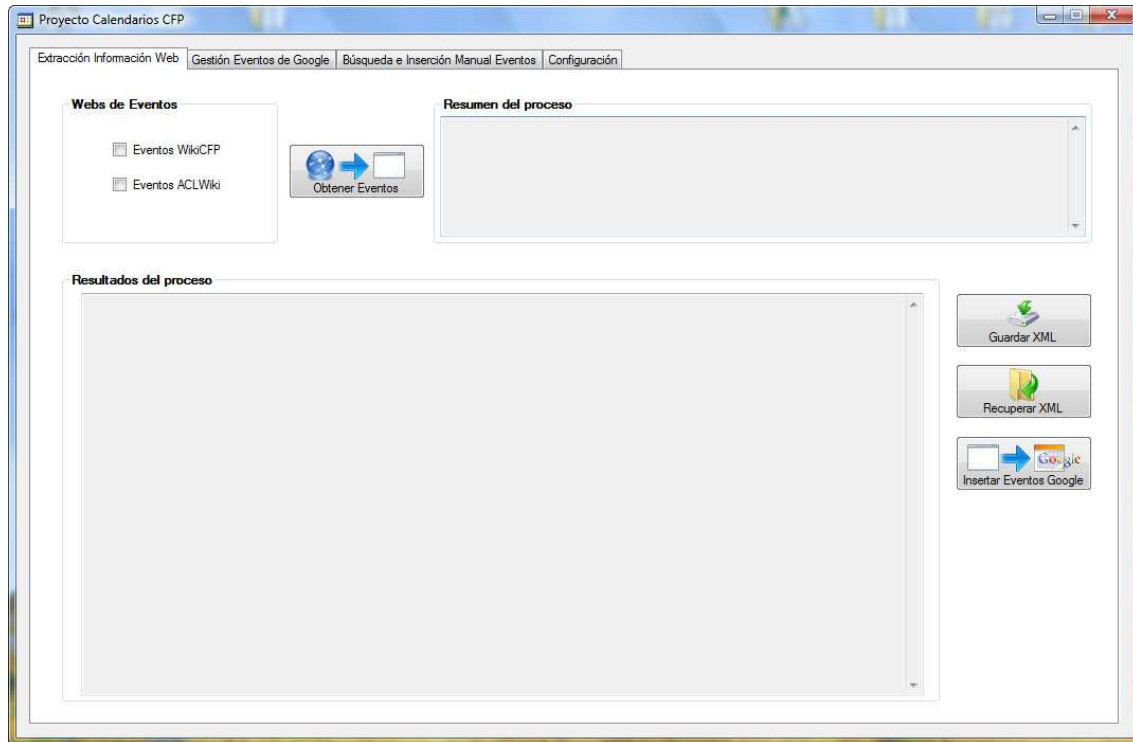
Se distinguen cuatro tipos de eventos en el sistema.

- Main Conference
- Deadline
- Abstract
- Workshops
- Full submission

Estos eventos son coincidentes con los eventos recuperados de las fuentes de extracción.

Con todos los datos tratados, se crean objetos de Google Calendar, que son las entradas de los eventos y se insertan. También se mostrará un resumen de este proceso.

### Interfaz de la pestaña de Extracción de Información Web (Figura 32)



**Figura 32.** Pestaña de extracción de información

## 5.2 PESTAÑA DE GESTIÓN DE EVENTOS DE CALENDARIO DE GOOGLE CALENDAR

En esta pestaña existe un botón a través del cual el sistema se conecta a Google Calendar, y obtiene los eventos accediendo al calendario creado previamente por la aplicación. Estos eventos se guardarán en un objeto *List* de .NET que contendrá objetos de tipos estructura de evento.

Los datos de los eventos recuperados se muestran en control DataGridView de .NET. Esta lista es ordenable por todas las columnas del evento.

Por otro lado, se irán guardando las localizaciones de los eventos en otro objeto *List*, para crear posteriormente con ellas la página HTML con código Javascript que muestra el mapa de Google Maps con las ubicaciones. Para incrustar la página HTML que contiene el mapa con los marcadores de las ubicaciones, se utiliza un objeto *WebBrowser*.

En esta pestaña se encuentran los botones de opción (*option buttons*). Hay cinco opciones:

1. Mostrar todos los eventos.
2. Mostrar los eventos para un día concreto.
3. Mostrar los eventos filtrando por su localización.  
Las localizaciones disponibles se guardan en otro objeto *List* cuando se recuperan los eventos. Estas ubicaciones se cargan en un control *ComboBox*.
4. Mostrar los eventos filtrando por su nombre.  
El filtrado es automático, buscando los eventos cuyo nombre empiece por los caracteres que se van insertando.
5. Mostrar los eventos filtrando por un rango de fechas.

Las búsquedas de estos eventos se realizan en el objeto *List* que contiene todos los eventos, sin necesidad de volver a conectar con Google Calendar.

El mapa se actualizará por cada recarga de eventos en la lista, de forma automática, mostrando en cada momento las ubicaciones de los eventos que se encuentren en la lista.

La lista muestra un menú contextual cuando se hace click con el botón derecho sobre un evento, permitiendo modificar o eliminar el evento seleccionado. En ambos casos después de realizar cada operación se conectará de nuevo con Google Calendar para modificar o eliminar la entrada del evento correspondiente en el calendario.

## Interfaz de la pestaña de Gestión de Eventos de Google (Figura 33)

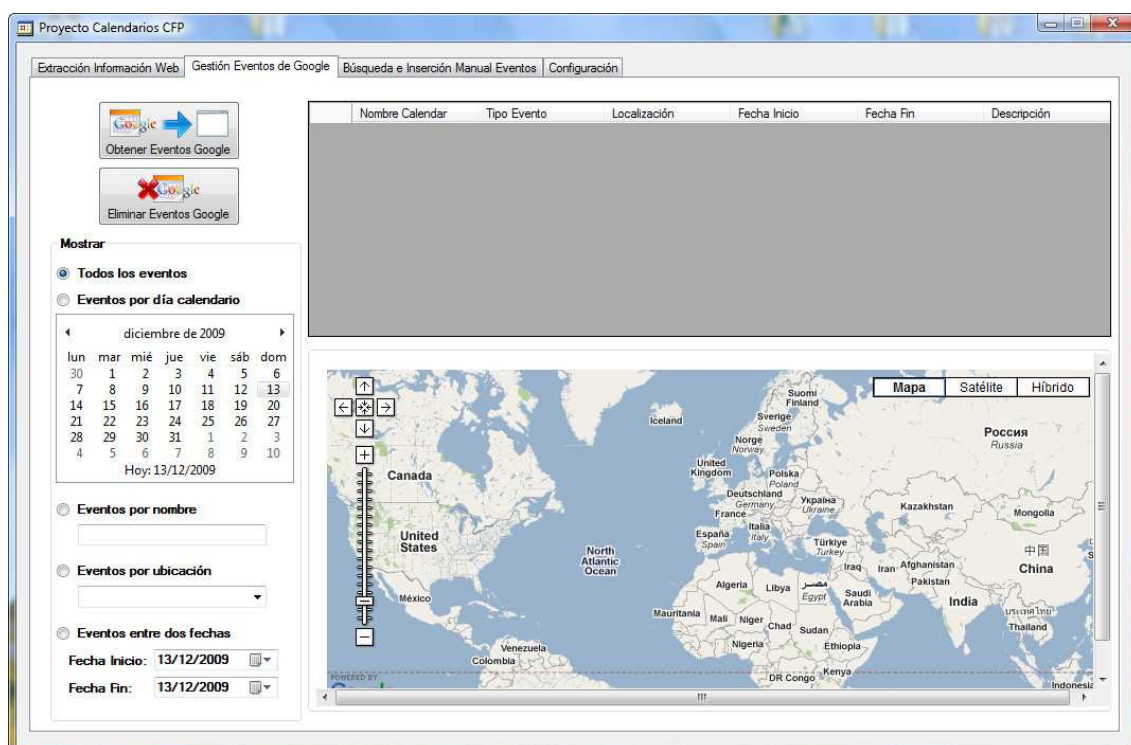


Figura 33. Pestaña de Gestión de Eventos de Google

### 5.3 PESTAÑA DE BÚSQUEDA E INSERCIÓN MANUAL DE EVENTOS

Esta pestaña incluye un objeto *WebBrowser* para mostrar la página HTML con código JavaScript que recupera el buscador de Google Search.

Contiene también el formulario para añadir un nuevo evento. Existen controles de datos para verificar que son correctos antes de que el evento se inserte en Google Calendar.

## Interfaz de la pestaña de Búsqueda e Inserción Manual de Eventos (Figura 34)

Proyecto Calendarios CFP

Extracción Información Web | Gestión Eventos de Google | **Búsqueda e Inserción Manual de Eventos** | Configuración

**Búsqueda en web**

Eventos Conferencias Científicas

con la tecnología de Google™

web

[Asesoría del Ámbito Científico: Conferencias científicas en el ...](#)  
29 May 2009 ... Conferencias científicas en el Ateneo de Almodóvar del Río ... Adjuntamos el cartel del evento realizado por Juan Rodríguez Ramos. ...  
[cepciencias.blogspot.com](#)

[Agenda de eventos científicos](#)  
Agenda de eventos científicos. Listando todos los eventos registrados ... Conferencia Mundial sobre Regulación de los Profesionales de la Salud ...  
[www.smu.org.uy](#)

[Informar desde las conferencias científicas - SciDev.Net](#)  
Será necesario estar dispuesto a asistir a todo tipo de conferencias científicas porque cada una tiene su utilidad. Pueden ser locales, regionales o ...  
[www.scidev.net](#)

[Conferencias de científicos - eventos.emagister.com](#)  
Conferencias de científicos. Las conferencias de científicos que no pueden faltar en tu agenda.  
[eventos.emagister.com](#)

1 2 3 4 5 6 7 8

**Inserción de Evento**

Nombre:

Nombre Calendar:

Descripción:

Localización:

Fecha Inicio: 13/12/2009

Fecha Fin: 13/12/2009

Tipo de Evento:

Figura 34. Pestaña de Búsqueda e Inserción Manual de Eventos

## 5.4 PESTAÑA DE CONFIGURACIÓN

Contiene los controles para recuperar y mostrar los datos configurables del sistema.

Estos datos se pueden modificar y guardar. Se agrupan de la siguiente forma:

- Datos de Acceso a Google Calendar
  - Usuario
  - Contraseña

La aplicación permite cambiar de usuario en tiempo de ejecución, guardando los datos de acceso del nuevo usuario en el fichero de configuración, y vaciando todos los controles, objetos y estructuras en memoria, para que posteriormente se puedan cargar los eventos del nuevo usuario.

- Opciones del proxy
- Fechas de recuperación de eventos del calendario creado por el sistema en Google Calendar.
- Nombre del calendario que se creará en Google Calendar. Este nombre es configurable. Si se cambia el nombre del calendario, también se vaciarán los controles, objetos y estructuras en memoria, para crear, ver o actualizar los eventos del nuevo calendario.

### Interfaz de la pestaña de Configuración (Figura 35)

*Figura 35.* Pestaña de Configuración

## 6 PRUEBAS

### 6.1 EJECUCIÓN DE PRUEBAS

A continuación se muestra una tabla (Tabla 165) que contiene los resultados de las pruebas realizadas para comprobar que el sistema cumple con los requisitos establecidos por el usuario.

Son necesarias para validar el sistema antes de la puesta en explotación.

ID	Nombre	Resultado
PR-001	Obtener eventos de una página Web	Satisfactorio
PR-002	Comprobación de eventos obtenidos	Satisfactorio
PR-003	Guardar el documento XML con los eventos en disco duro	Satisfactorio
PR-004	Recuperar el fichero con los eventos del disco duro	Satisfactorio
PR-005	Inserción de eventos en Google Calendar	Satisfactorio
PR-006	Recuperar y mostrar eventos de Google Calendar	Satisfactorio
PR-007	Pruebas de filtros de eventos del sistema	Satisfactorio
PR-008	Modificar los datos de un evento de la lista de eventos	Satisfactorio
PR-009	Eliminar un evento de la lista de eventos	Satisfactorio
PR-010	Eliminar todos los eventos del sistema	Satisfactorio
PR-011	Buscar páginas con el buscador de Google Search	Satisfactorio
PR-012	Añadir evento de forma manual	Satisfactorio
PR-013	Cambiar los datos de acceso del usuario en pestaña configuración	Satisfactorio
PR-014	Cambiar fechas de recuperación de eventos en pestaña configuración	Satisfactorio
PR-015	Probar opciones de proxy en pestaña configuración	Satisfactorio
PR-016	Introducir nombre de Calendario en pestaña configuración	Satisfactorio

**Tabla 165.** Resultados de las pruebas de validación del sistema

## **6.2 ACCIONES CORRECTIVAS**

No aplica.



## 7 CONCLUSIONES

Tras varios meses de trabajo, de pruebas y de investigación, se ha podido concluir con éxito este proyecto.

Durante el desarrollo del mismo se ha tenido oportunidad de conocer cómo está funcionando actualmente el entorno Web así como las complicaciones que existen para poder acceder a los contenidos, apreciándose notablemente la diferencia de cómo se ha realizado y cómo se podría realizar si el entorno fuese diferente. Si se tiende a la Web semántica, el trabajo se facilitaría enormemente y daría lugar a aplicaciones comerciales consistentes, que se basarían en otras aplicaciones Web. Esto sería posible a través de la importancia de seguir un formato estándar para mostrar los contenidos en las páginas publicadas, lo que daría lugar a sistemas de recopilación de información más robustos, con apenas unas pocas líneas de código.

Desafortunadamente ese no es el caso para las páginas de conferencias científicas, donde la forma de estructurar la información de cada página es totalmente diferente de las demás. Es por ello que se ha procedido a realizar el parseo de la información de las páginas de la aplicación de forma particular para cada una de ellas. Esto está supeditado a conocer previamente como está estructurado el código fuente de cada una, y a las limitaciones de funcionamiento si se cambia la estructura de la página, siendo imposible conocer de forma estándar el contenido referente a los eventos que se encuentran en ellas.

Ha resultado interesante, a la par que complicado, el manejo del API de Datos de Google Calendar. El API es poco intuitivo y se ha necesitado bastante tiempo de estudio de estas librerías para poder integrarlas en la aplicación desarrollada. En la documentación del API se facilitan ejemplos que ayudan “parcialmente” a integrar estas funcionalidades en el sistema. En su integración aparecen varios errores. Algunos de ellos han requerido bastante tiempo de investigación, y se ha comprobado que son errores comunes y con poca documentación para resolverlos. De estos errores existen muchas dudas que los desarrolladores exponen en la Web y poca información al respecto. Como parte positiva, es una tecnología actual, novedosa, y que aporta unas

posibilidades de integrar, gestionar y sincronizar calendarios de Google Calendar en una aplicación de escritorio, accediendo a través la cuenta de un usuario, utilizando como almacén de datos el gestor de Google. Como parte negativa, comentar que su integración requiere una buena parte de investigación y que el acceso a Google Calendar a través de los métodos del API es bastante lento, y requiere de una conexión con un ancho de banda amplio para que la ejecución no se ralentice demasiado.

Se ha probado también el funcionamiento de las API de AJAX de Google para Google Maps y Google Search, con la integración del mapa de localización de eventos y del buscador de información en páginas Web, respectivamente. La integración del mapa a través del API de Google Maps se complica cuando se quieren ubicar varios puntos de localización de forma dinámica, sin conocer previamente sus coordenadas. Las llamadas que se hacen a los métodos de éste API son asíncronas, y hacen que se complique. El cálculo interno que hace el propio Google Maps aplicando el método para calcular las coordenadas de una ubicación es lento, y cuando se tiene que hacer para varias ubicaciones, como es el caso, lo es aún más. Como parte positiva la integración del mapa en la aplicación hace que sea más funcional y más vistosa, permitiendo ubicar geográficamente los eventos de las conferencias en el mapa, rápidamente.

En cuanto a la integración del buscador de páginas Web, a diferencia de las otras APIs, se realiza de forma sencilla, y su uso está muy documentado.

## 8 LÍNEAS FUTURAS

A continuación se comentan algunas ampliaciones que se pueden realizar sobre la aplicación en trabajos futuros:

- a) Reconocer los distintos formatos de hora dentro de la fecha del deadline. Por ejemplo: GMT, AM, PM...

Actualmente la fecha del deadline se muestra en un formato estándar: “dd/mm/yyyy”, sin especificar la hora. Esto se puede ampliar de manera que se recuperen también otros tipos de formato y se puedan introducir horas.

- b) Clasificación por continente en el mapa.

Sería interesante investigar la forma de poder localizar los continentes donde se realizan los eventos. En las páginas donde se ha realizado la extracción automática de eventos, el lugar de celebración indica el país pero no el continente. Sería preciso buscar un sistema donde almacenar los continentes, con los países se que encuentran dentro de los mismos. De forma automática se debería localizar el país dentro de la dirección del evento recuperado y consultando dentro del sistema de información anterior, recuperar el continente en el que se celebrara dicho evento.

Habría que ubicar posteriormente el continente en el mapa con sus marcadores, y en el evento clic de cada continente mostrar los eventos que tengan lugar en el mismo.

- c) Combinación de filtros. Por ejemplo:
  - Filtrado de eventos por un día del calendario y un lugar de celebración.
  - Filtrado de eventos por un lugar de celebración y una fecha o un rango de fechas.
- d) Integrar colores en función del tipo de evento.

Cuando la aplicación recupera los eventos del calendario de Google Calendar, actualmente hay un método que distingue el tipo de evento antes de insertarlo en la lista. Se podría incluir un nuevo método que cambie el color de la fuente o el color de fondo para cada la fila de la lista en función del tipo de ese evento que se vaya a insertar.

- e) Ampliar la funcionalidad del sistema software desarrollado para hacerlo más genérico, de manera que sirva para otras aplicaciones de organización y gestión de otros tipos de eventos (agenda de trabajo, planificación de viajes, etc).
- f) Evaluación de la precisión del módulo de extracción de información a la hora de interpretar eventos. Mostrar estadísticas sobre la precisión de los resultados obtenidos.
- g) Si en algún momento se implanta la Web Semántica en el ámbito de las conferencias científicas mediante el uso de microformatos, adaptar la aplicación de manera que el módulo de extracción de información sea capaz de interpretarlos.

## 9 LISTA DE ACRÓNIMOS

AMD	Advanced Micro Devices
API	Application Programming Interface
AJAX	Asynchronous JavaScript And XML
CERN	European Organization for Nuclear Research
CFA	Call For Abstract
CFP	Call For Paper
DOM	Document Object Model
DTD	Document Type Definition
ECMA	European Computer Manufacturers Association
FAQ	Frequently Asked Question
GPS	Global Positioning System
HTML	HyperText Markup Language
IA	Itanium Architecture
IDE	Integrated Development Environment
IETF	Internet Engineering Task Force
iFrame	inline Frame
ISO	International Organization for Standardization
JSON	JavaScript Object Notation
MSDN	MicroSoft Developer Network
PCO	Proffesional Conference Organizer
PLN	Procesamiento del Lenguaje Natural
POS	Part-Of-Speech
RDFa	Resource Description Framework attributes
RFC	Request For Comments
RSS	Really Simple Syndication
SDK	Software development kit
SGML	Standard Generalized Markup Language
VML	Vector Markup Language
URL	Uniform Resource Locator
WHATWG	Web Hypertext Application Technology Working Group
W3C	World Wide Web Consortium

XFN	XHTML Friends Network
XHTML	eXtensible Hypertext Markup Language
XMDP	XHTML Meta Data Profiles
XML	eXtensible Markup Language
XOXO	eXtensible Open XHTML Outlines
XSD	XML Schema Definition
XSL	eXtensible Stylesheet Language

## 10 BIBLIOGRAFÍA

### 10.1 MATERIAL REFERENCIADO

- [1]. API de datos de Google. Incluye el API para Google Calendar. Disponible en:  
[<http://code.google.com/intl/es/apis/gdata/>]
- [2]. API AJAX de Google. Incluye las APIs de Google Maps y de Google para búsquedas (Search). Disponible en: [<http://code.google.com/intl/es/apis/ajax/>]
- [3]. Microformats.org, comunidad on-line que se encarga de mantener y especificar la mayoría de los microformatos existentes. [<http://www.microformatos.org>]
- Contiene un artículo publicado en Marzo del 2006 que recoge la opinión de Bill Gates sobre los microformatos. Disponible en:  
[<http://microformats.org/2006/03/20/bill-gates-at-mix06-we-need-microformats>]
- [4]. Web Standards Project. [<http://www.webstandards.org/>]
- [5]. Proyecto ANDES:  
Jussi Myllymaki. *Effective Web data extraction with standard XML technologies*. Computer Networks, Volume 39, Number 5, 5 August 2002 , pp. 635-644(10).  
Jussi Myllymaki, Jared Jackson. *Automatically extract information with HTML, XML, and Java*. 2001.
- [6]. Proyecto ARIADNE:  
C. A. Knoblock, S. Minton, J. L. Ambite, N. Ashish, I. Muslea, A. G. Philpot, and S. Tejada. *The ariadne approach to Web-based information integration*. International the Journal on Cooperative Information Systems (IJCIS), Special Issue on Intelligent Information Agents: Theory and Applications, 10(1):145--169, 2001.
- [7]. *Web 2.0, Web 3.0 y Web 4.0*. Beatriz Helena Melo Herrera. Universidad ICESI. Mayo 2009. Disponible en:  
[<http://www.icesi.edu.co/blogs/egatic/2009/05/21/web-20-web-30-y-web-40/>]
- [8]. ISO8601:2004(E) *Data elements and interchange formats – Information interchange – Representation of dates and times. Third Edition*. 2004.

- [9]. *Análisis, Diseño y Mantenimiento del Software - Ciclo de vida del software*. José Ramón Álvarez Sánchez y Manuel Arias Calleja. Dpto. de Inteligencia Artificial - ETSI Informática – UNED. Diciembre 2002.
- [10]. Extensible Markup Language (XML) 1.0 (5th Edition) W3C Recommendation. Noviembre 2008. Disponible en:  
[<http://www.w3.org/TR/REC-xml/>]
- [11]. WikiCFP - A Wiki for Calls for Papers. Disponible en:  
[<http://www.wikicfp.com/cfp/call?conference=NLP&skip=1>]
- [12]. ACL Wiki for Computational Linguistics. Wiki de la ACL (Association for Computational Linguistics). Disponible en:  
[[http://aclweb.org/aclwiki/index.php?title=Conferences\\_and\\_workshops](http://aclweb.org/aclwiki/index.php?title=Conferences_and_workshops)]
- [13]. Dmoz Open Directory Project – Directorio Web con información sobre conferencias Disponible en:  
[[http://www.dmoz.org/Computers/Computer\\_Science/Conferences](http://www.dmoz.org/Computers/Computer_Science/Conferences)]
- [14]. Association for Computations Linguistics. [<http://www.aclweb.org/>]

## 10.2 MATERIAL DE APOYO

Para concluir la sección de referencias, incluiremos una serie de materiales que, a pesar de no haber sido directamente referenciados en la memoria, han servido de apoyo para la elaboración del proyecto.

- David Flanagan; Paula Ferguson (2002). *JavaScript: The Definitive Guide*, 4th Edition, O'Reilly & Associates. ISBN 0-596-00048-0.
- David Flanagan. *JavaScript: The Definitive Guide*, 5th Edition (2006). O'Reilly Media, Inc. ISBN 0-596-10199
- Rainer Kolbeck (1997). *El gran libro de Javascript*. Editorial Marcombo. ISBN 84-267-1113-8.
- Mario G. Piattini, José A. Calvo, Joaquín Cervera, Luis Fernández (2003). *Aplicaciones Informáticas de Gestión. Una perspectiva de Ingeniería del Software*. Editorial Ra-Ma. ISBN 84-7897-587-X
- Fco. Javier Ceballos (2006). *Enciclopedia de Microsoft Visual C#*. Editorial Ra-Ma. ISBN 84-7897-707-4



- J.A Zabalegui (2007). *Recuperación y Organización de la información*. Disponible en: [<http://extraccioninformacion.iespana.es/>]
- R. Olías Beltrán. Extracción y Recuperación de la Información I.UC3M. Disponible en: [<http://recupera1patrones.50webs.com/>]
- Salvador Perez Crespo (2007). *Cómo será la Web 3.0*. Telefónica I+D para Telefónica España. Disponible en: [<http://sociedadinformacion.fundacion.telefonica.com/>]
- Lenguajes de programación usados en Internet y la World Wide Web (WWW). Disponible en: [[http://www.webdevelopersnotes.com/basics/languages\\_on\\_the\\_internet.php3](http://www.webdevelopersnotes.com/basics/languages_on_the_internet.php3)]
- Breve historia del HTML. Disponible en: [[http://librosweb.es/xhtml/capitulo1/breve\\_historia\\_de\\_html.html](http://librosweb.es/xhtml/capitulo1/breve_historia_de_html.html)]
- APIs de Google. Código, documentación, etc.: [<http://code.google.com/intl/es/more/>]
- Wikipedia,
  - Microformatos: [<http://es.wikipedia.org/wiki/Microformato>]
  - Conceptos relacionados con los CFPs: [[http://en.wikipedia.org/wiki/Academic\\_conference](http://en.wikipedia.org/wiki/Academic_conference)]  
[[http://en.wikipedia.org/wiki/Call\\_for\\_papers](http://en.wikipedia.org/wiki/Call_for_papers)]

## **11 ANEXOS**

### **ANEXO 1 – MANUAL DE USUARIO**

A continuación se procede a explicar el funcionamiento de la aplicación para que cualquier usuario que desconozca el sistema, sea capaz de manejarlo.

El manual se va a estructurar en función de las operaciones que se pueden realizar en las distintas pestañas:

- a) Pestaña de Configuración.
- b) Pestaña de Extracción de Información Web.
- c) Pestaña de Gestión de Eventos de Google.
- d) Pestaña de Búsqueda e Inserción Manual de Eventos.

#### **A. Pestaña de Configuración**

Se va a comenzar a explicar la pestaña de configuración puesto que los datos de configuración influyen en el resto de operaciones del sistema:

En la siguiente imagen (Figura 36), se muestra la pestaña de configuración:

**Figura 36.** Pestaña de datos de configuración

Cuando se comienza a usar la aplicación, el sistema carga unos datos por defecto, que posteriormente, pueden ser modificados. Estos datos se agrupan en cuatro bloques:

- Datos de conexión con Google Calendar.

Estos datos se corresponden con la cuenta que el usuario debe haber creado previamente en el gestor de correo de Google: gmail. La aplicación se conectará con el servicio de Google Calendar, a través de estos datos.

- Datos de configuración de proxy.

Si la conexión a Internet requiere la configuración de un proxy, hay que introducir el servidor y el número de puerto en las cajas de texto correspondientes y desmarcar el checkbox *No utilizar proxy*. Si por el contrario, no es necesario configurar un proxy, el checkbox *No utilizar proxy* tiene que estar marcado.

- Fechas de obtención de eventos de Google Calendar.

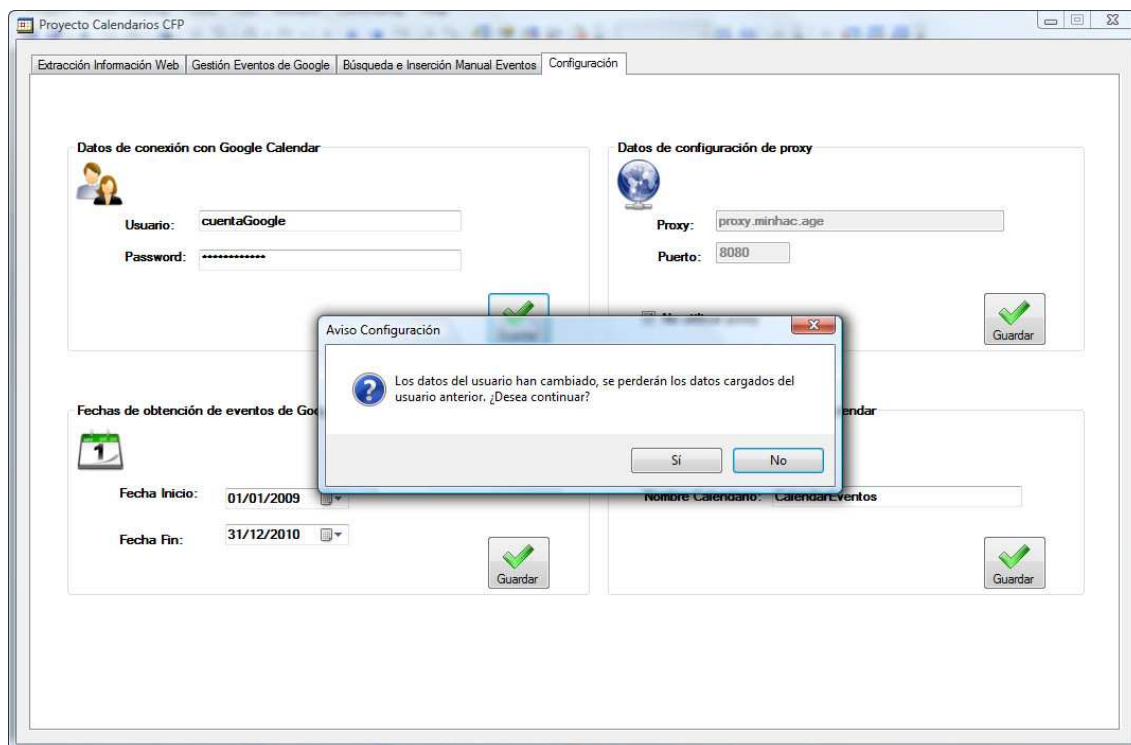
Para obtener los eventos de un calendario de Google Calendar, hay que especificar la fecha de inicio y la fecha fin de obtención de eventos.

- Calendario de Eventos de Google Calendar.

En la caja de texto de *Nombre de calendario*, es donde se especifica el nombre del calendario que se creará dentro de Google Calendar. Sobre este calendario se realizarán las gestiones de los eventos.

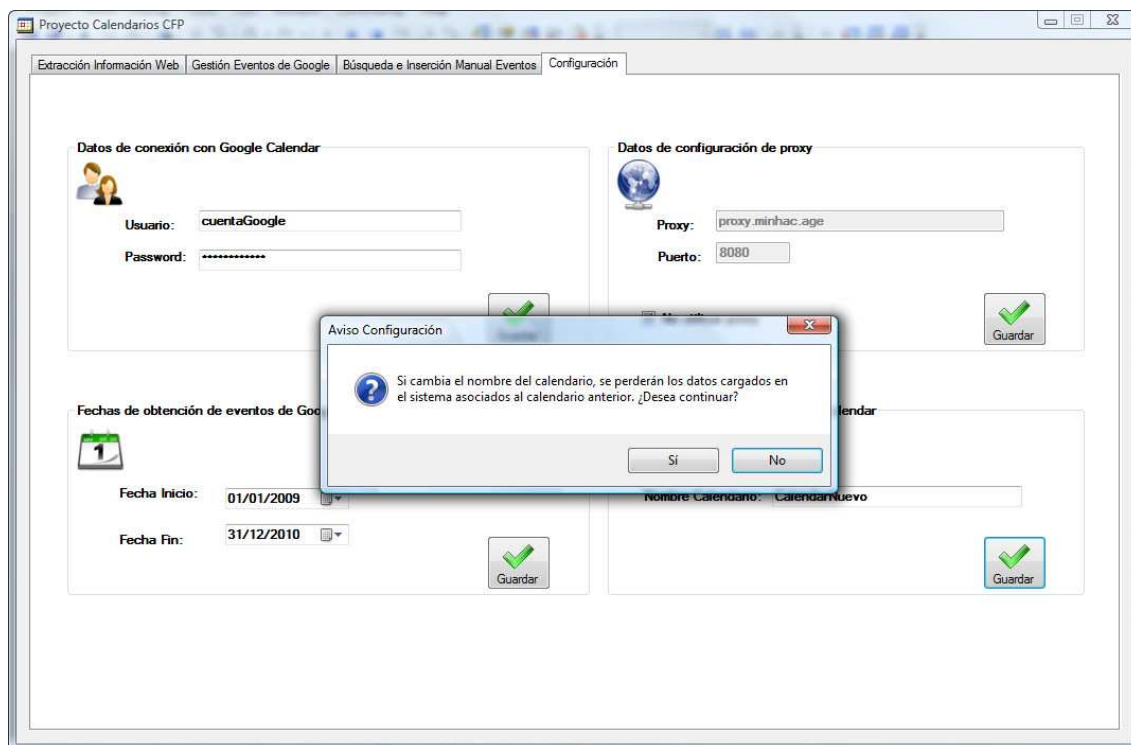
El botón *Guardar* de cada bloque, guarda los datos modificados en el fichero de configuración.

Cuando se modifican los datos del usuario, se reinicia todo el sistema para que se pueda comenzar a operar con los datos del nuevo usuario. Estos datos se pueden modificar en tiempo de ejecución, lo que implica que si se había operado previamente con otro usuario, la aplicación muestra un mensaje avisando de que el usuario ha cambiado, y se perderán los datos del usuario anterior, como se muestra en la Figura 37:



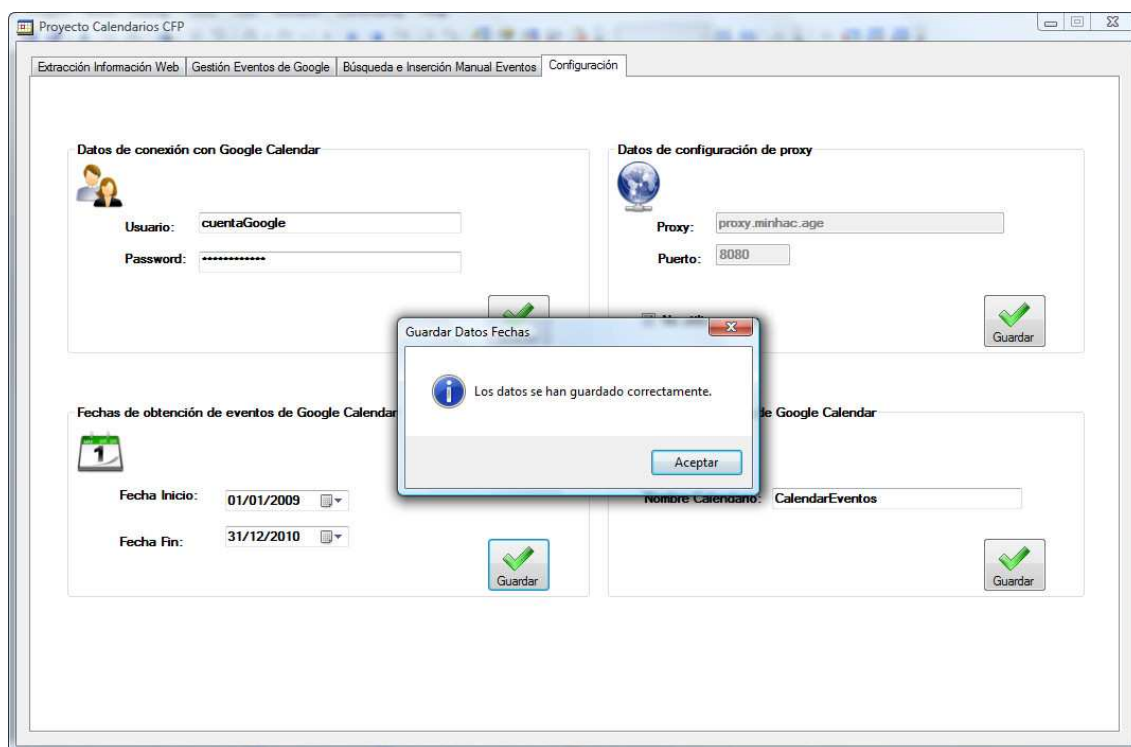
**Figura 37.** Confirmación del cambio en los datos de conexión del usuario

Algo similar sucede cuando se cambia el nombre del Calendario. El sistema se reiniciará, preparándose para gestionar los eventos con el nuevo calendario, mostrando también un mensaje de aviso, como se puede ver en la Figura 38:



**Figura 38.** Confirmación al cambiar el nombre del calendario

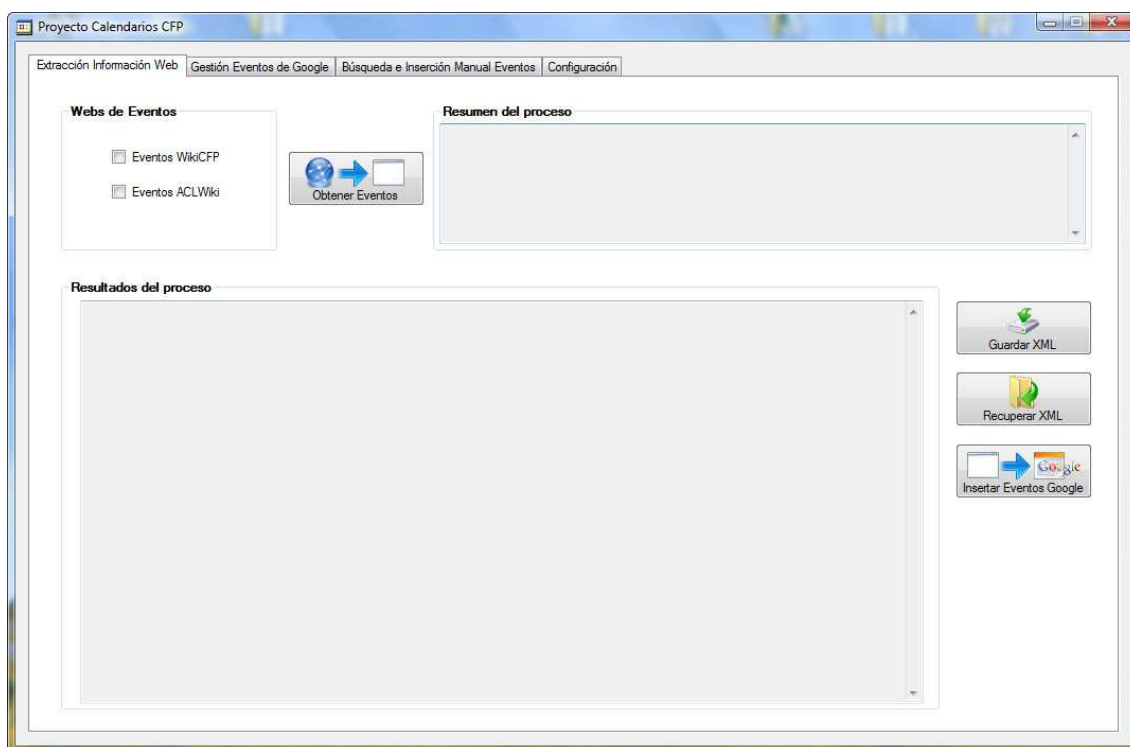
Cuando se guardan los datos, el sistema muestra un mensaje de aviso, indicando que se ha realizado correctamente. Esto se puede ver en la Figura 39:



*Figura 39.* Aviso de datos guardados correctamente

## B. Pestaña de Extracción de Información Web

A continuación se procede a explicar el funcionamiento de la pestaña de Extracción de Información Web (ver Figura 40):



**Figura 40.** Pestaña de Extracción de Información Web

En esta pestaña se pueden realizar las siguientes operaciones:

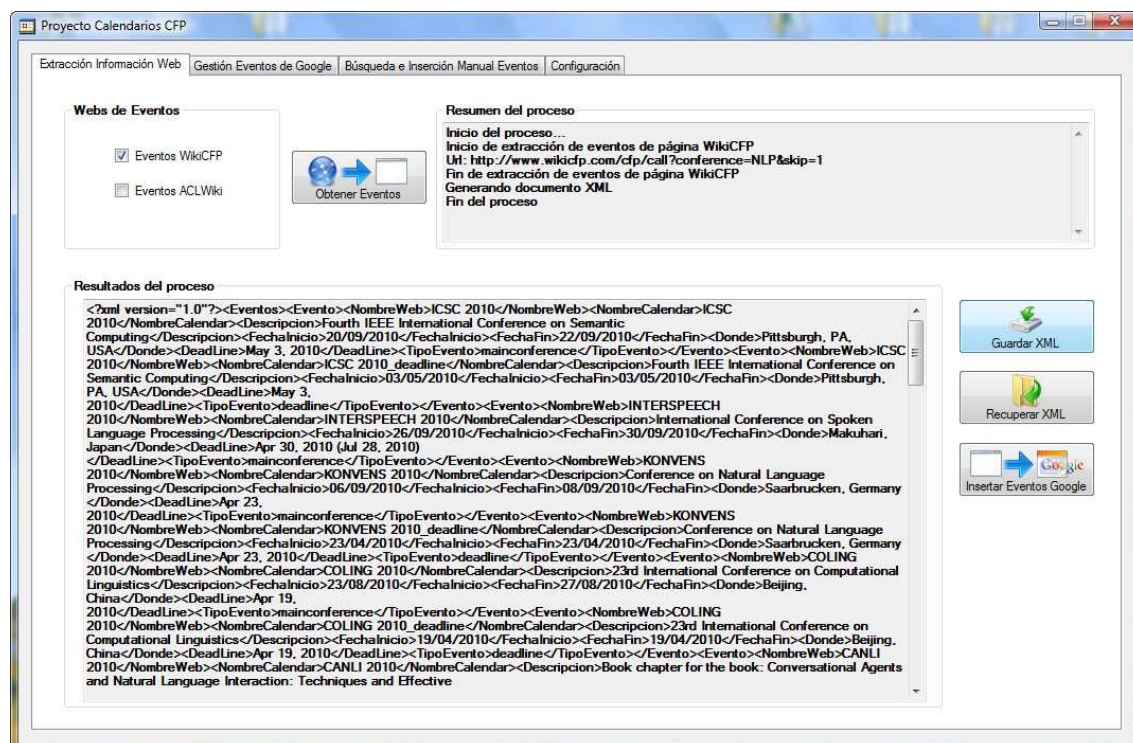
- 1) Obtener eventos de páginas Web.
- 2) Guardar documento XML en un fichero.
- 3) Recuperar documento XML de un fichero.
- 4) Inserción de eventos recuperados a un calendario de Google Calendar.

### **1) Obtener eventos de páginas Web**

Para comenzar con la extracción de información, se debe marcar al menos uno de los *checkboxes* de la sección de *Webs de Eventos*, correspondiéndose cada uno con un origen de datos susceptibles de extracción. A continuación se debe hacer click sobre el botón de *Obtener Eventos*.

Una vez que los eventos son obtenidos, en el cuadro de texto de *Resultados del Proceso*, se mostrarán los datos de los eventos obtenidos en formato XML. El cuadro de texto de *Resumen del Proceso* mostrará información sobre el proceso de extracción. Esto se

puede apreciar en la siguiente imagen (Figura 41), donde se ha seleccionado la Web *Eventos WikiCFP*:

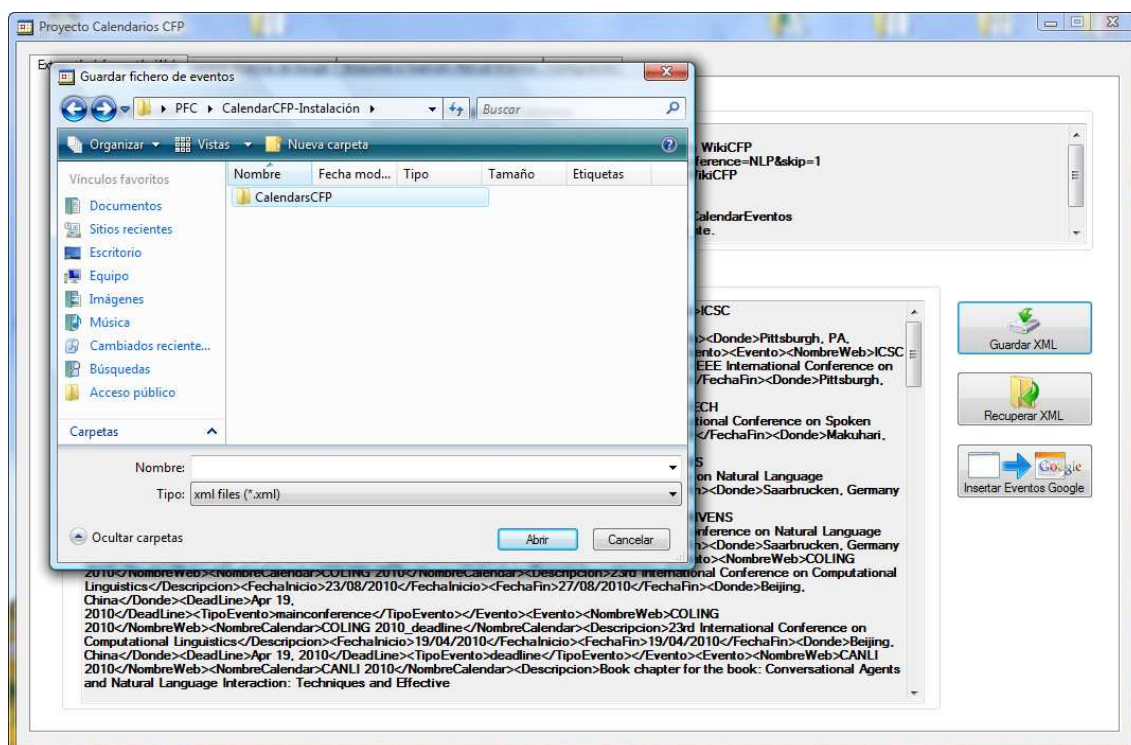


**Figura 41.** Información mostrada después de Obtener Eventos de un origen de datos

## 2) Guardar documento XML en un fichero

Mediante el botón *Guardar XML*, el sistema permite almacenar los datos de eventos obtenidos en un fichero .xml dentro del disco duro. Para ello se dispone de un cuadro de diálogo, como se puede ver en la Figura 42:





**Figura 42.** Guardar documento XML en un fichero

### 3) Recuperar documento XML de un fichero

De la misma forma, el sistema permite recuperar el fichero o ficheros guardados anteriormente, haciendo click en el botón *Recuperar XML*. También para esta operación se muestra un cuadro de diálogo (ver Figura 43):

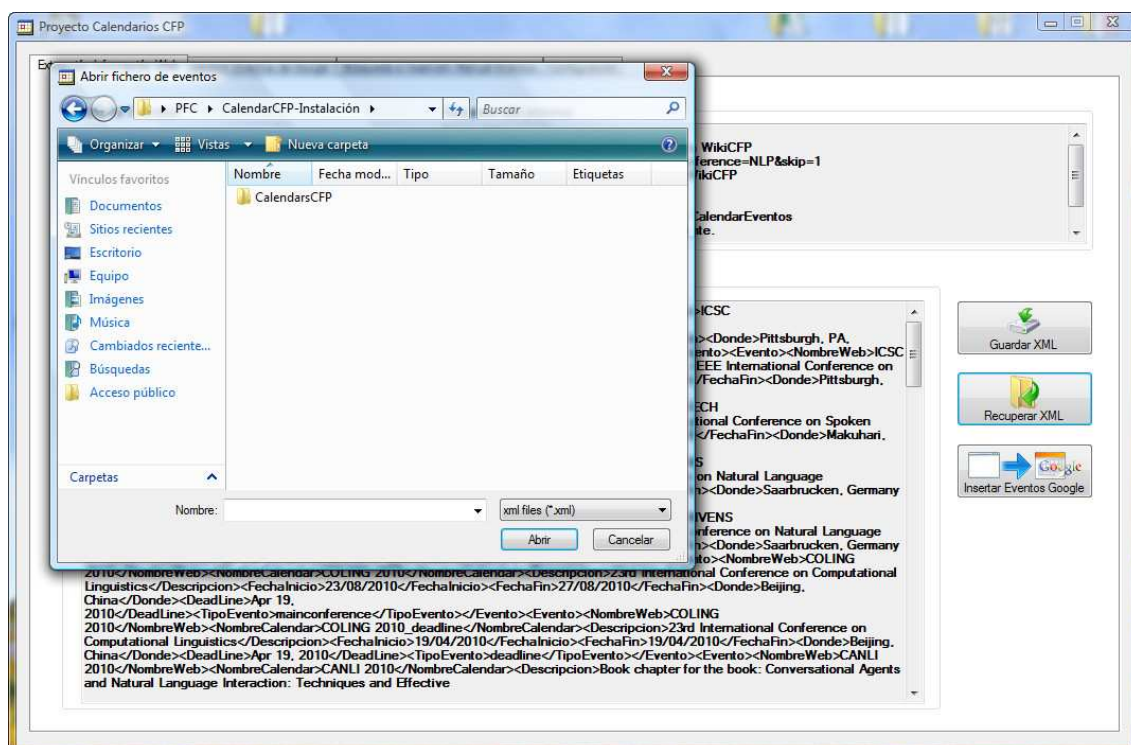


Figura 43. Recuperar documento XML de un fichero

Cuando se recuperan los eventos del fichero de eventos, el sistema actualiza sus datos con los nuevos eventos. Esta operación quedará almacenada en la caja de texto de *Resumen del proceso*.

#### 4) Inserción de eventos recuperados a un calendario de Google Calendar

Otra operación importante que se puede realizar, es la inserción de los eventos recuperados de la extracción de información a un calendario de Google Calendar. El nombre de este calendario es el que se especifica en la parte de datos de configuración (como se ha visto anteriormente). Para llevar a cabo esta operación el usuario debe hacer click sobre el botón *Insertar Eventos Google*. La operación una vez más quedará almacenada en la caja de texto de *Resumen del proceso*.

### C. Pestaña de Gestión de Eventos de Google

A continuación se explica el funcionamiento de la pestaña de Gestión de Eventos de Google (Figura 44):

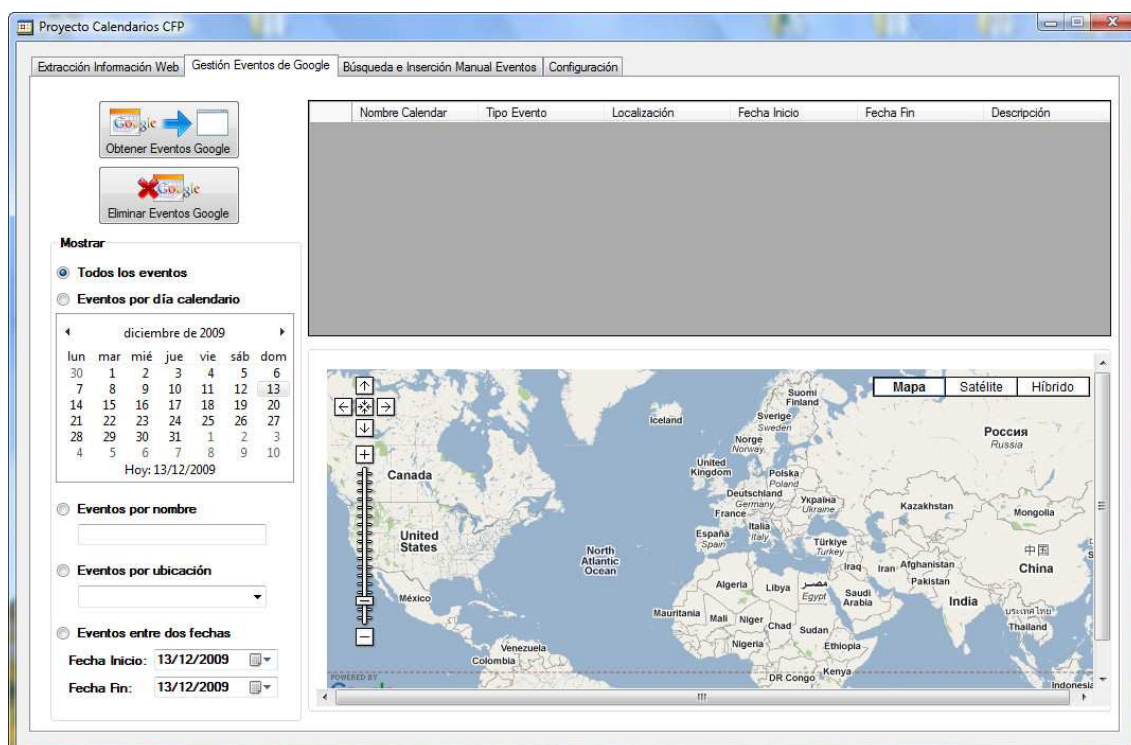


Figura 44. Pestaña de Eventos de Google

En esta pestaña se pueden realizar las siguientes operaciones:

- 1) Obtener Eventos del calendario de Google Calendar.
- 2) Eliminar todos los eventos del calendario de Google Calendar.
- 3) Eliminar el evento seleccionado.
- 4) Modificar el evento seleccionado.
- 5) Mostrar evento en función del filtro seleccionado.

### 1) Obtener Eventos del calendario de Google Calendar

Obtener Eventos del calendario creado en Google. Para ello se hace click sobre el botón *Obtener Eventos Google*.

Cuando los eventos se han recuperado del calendario de Google Calendar, aparece marcada la opción de *Mostrar Todos los eventos*, los eventos figuran cargados en la lista, y sus ubicaciones en el mapa. Esto se puede ver en la Figura 45:

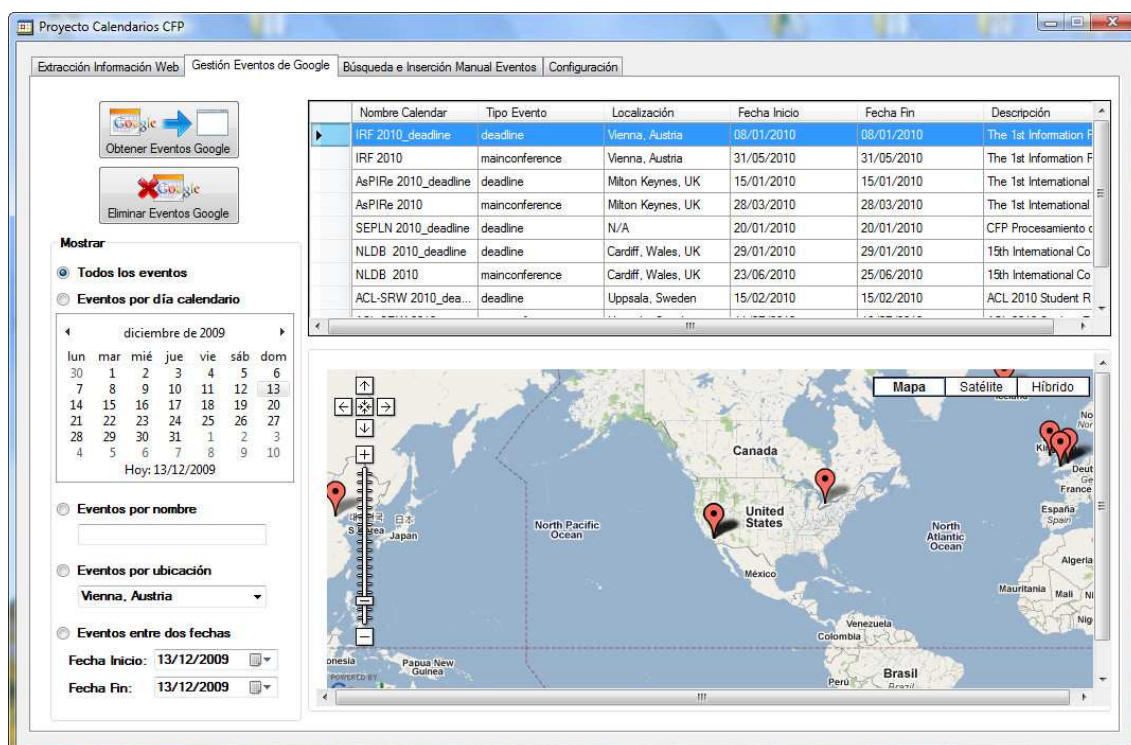
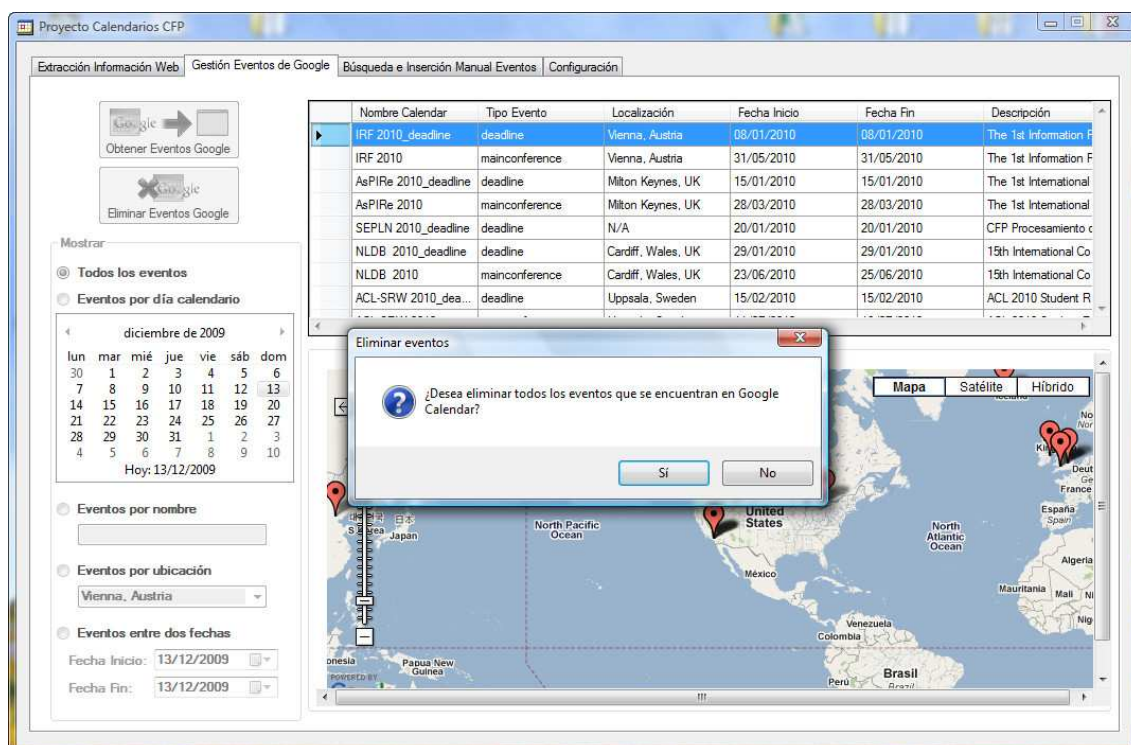


Figura 45. Pestaña de Eventos de Google después de obtener eventos

## 2) Eliminar todos los eventos del calendario de Google Calendar

Una vez que los eventos están cargados en el sistema, se puede Eliminar todos los eventos, haciendo click sobre el botón *Eliminar Eventos Google*. Después de esto, aparecerá un mensaje de confirmación de borrado, como se puede ver en la Figura 46:



**Figura 46.** Aviso de confirmación de borrado de todos los eventos

Una vez que el usuario seleccione *Sí*, se eliminan todos los eventos de la aplicación y del calendario de Google Calendar.

### 3) Eliminar el evento seleccionado

Si se selecciona un evento en la lista, cuando se hace click con el botón derecho, aparece un menú contextual permitiendo eliminar o modificar un evento (ver Figura 47):



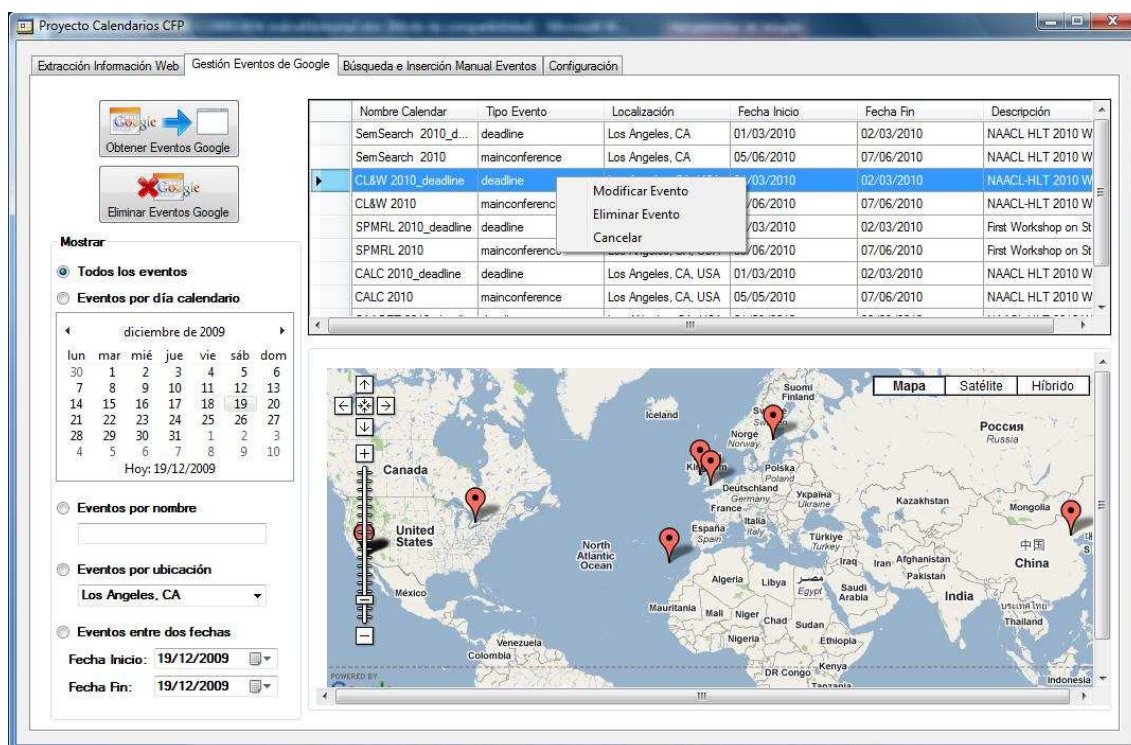


Figura 47. Menú contextual asociado a un evento seleccionado en la lista

Si se hace click en *Eliminar Evento*, aparece un mensaje de confirmación de borrado (Figura 48):

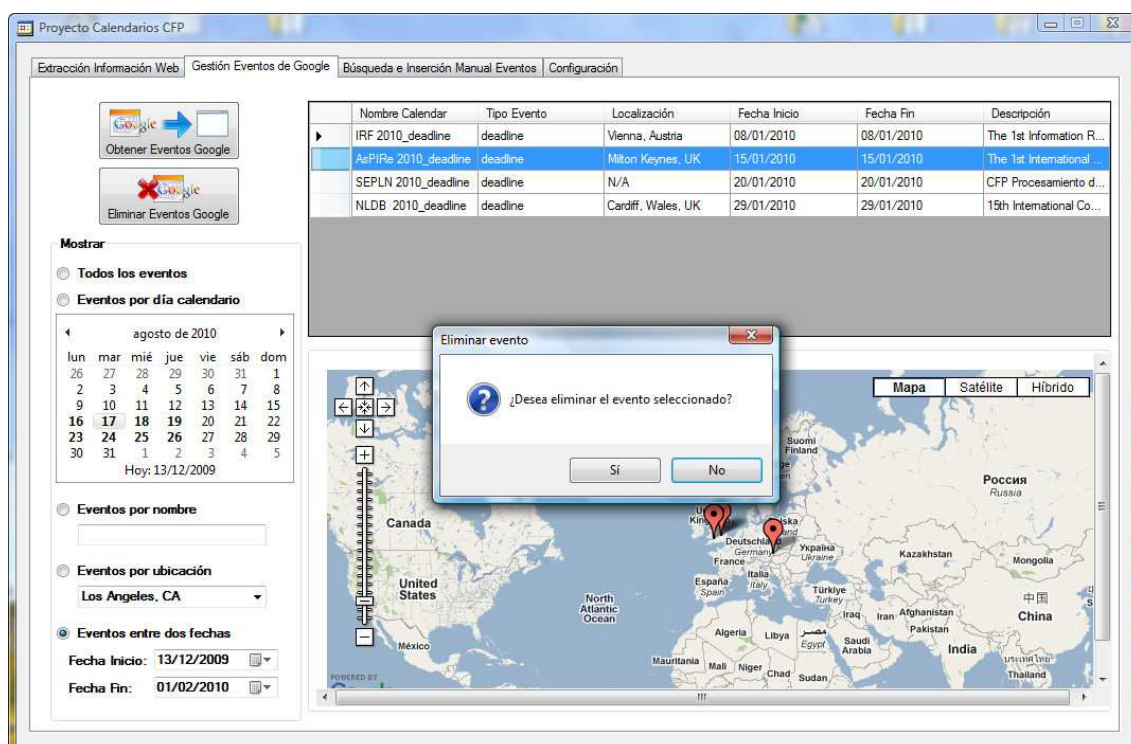
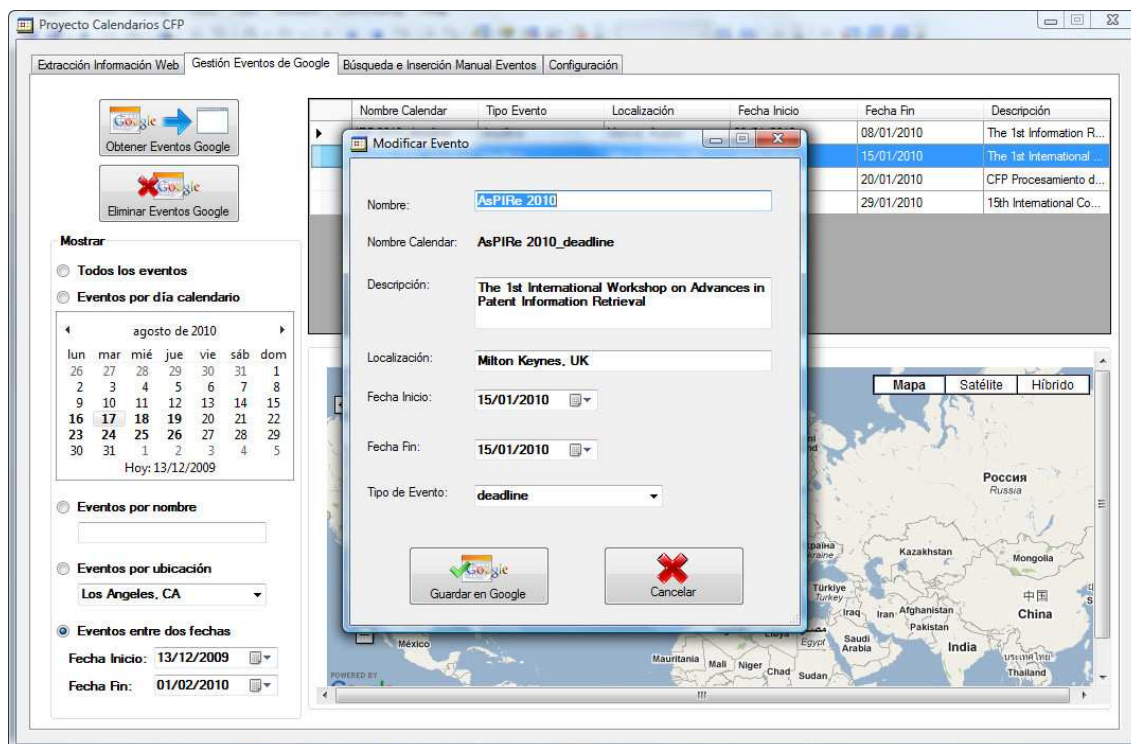


Figura 48. Confirmación de borrado del evento seleccionado

Cuando el usuario elige *Si*, este evento se elimina automáticamente de la lista y del calendario de Google Calendar.

#### 4) Modificar el evento seleccionado

Si se hace click en la opción de *Modificar Evento* del menú contextual de la Figura 47, se muestra un formulario de edición del evento con todos sus datos cargados (ver Figura 49):



**Figura 49.** Formulario de edición de un evento

Una vez modificados los datos deseados del formulario se hace click en el botón *Guardar en Google*, y automáticamente los datos correspondientes aparecerán actualizados en la pestaña de *Gestión de Eventos de Google*.

#### 5) Mostrar evento en función del filtro seleccionado

Las opciones de Filtrado de eventos, están agrupadas en el contenedor *Mostrar*. Cuando se hace click sobre cada opción se cargan en la lista los eventos filtrados, y en el mapa las localizaciones geográficas correspondientes a los mismos.

## Las opciones son:

- **Todos los eventos.** Es la opción que aparece seleccionada por defecto cuando se recuperan los eventos del calendario de Google Calendar. Muestra todos los eventos existentes (ver Figura 50).

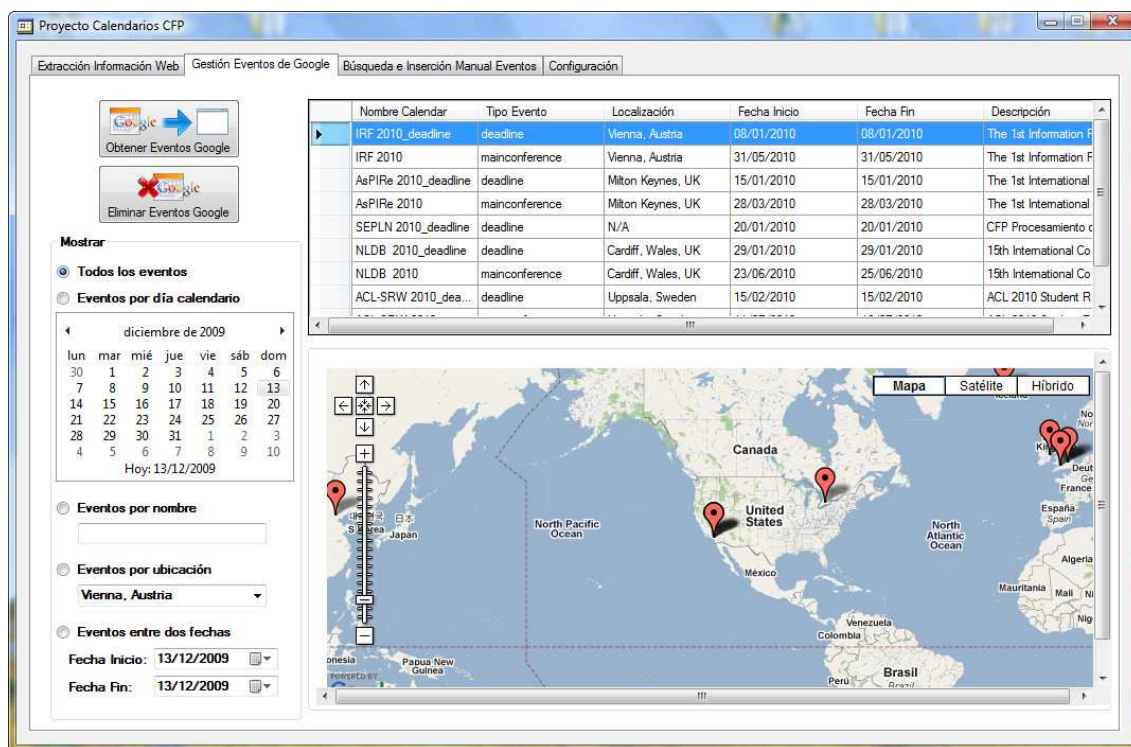


Figura 50. Mostrar todos los eventos

- **Eventos por día calendario.** Muestra los datos de los eventos para el día del calendario seleccionado (ver Figura 51).



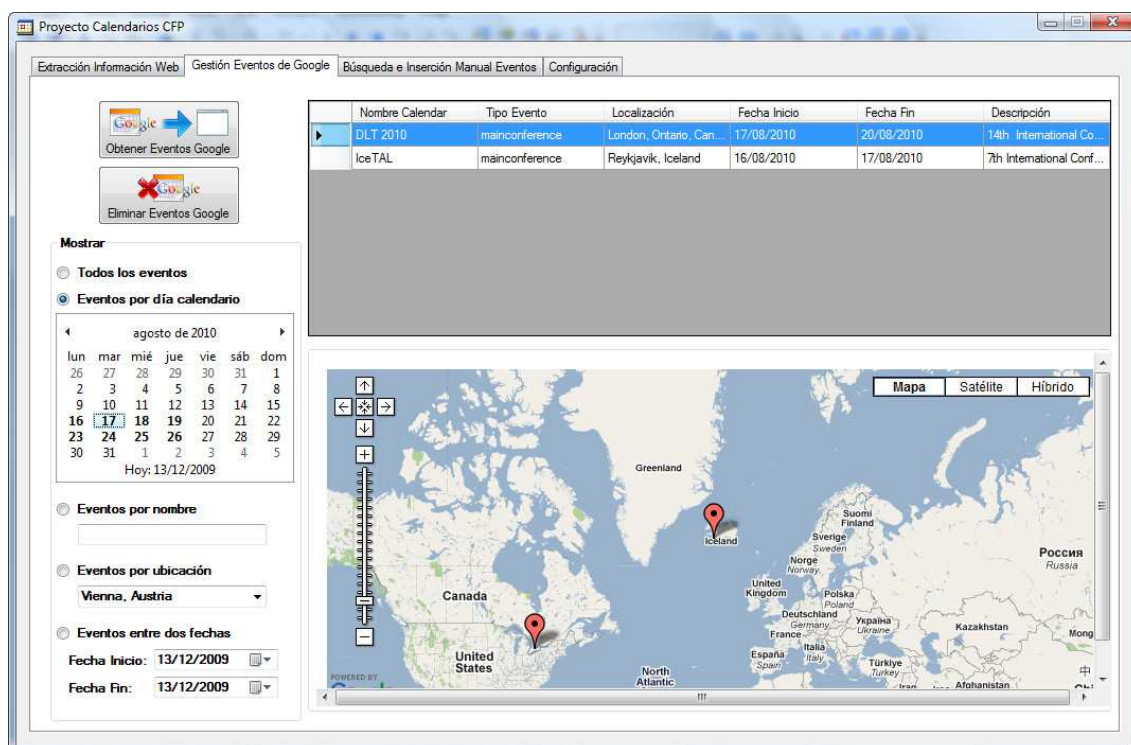


Figura 51. Mostrar los eventos para el día seleccionado en el calendario

- **Eventos por nombre.** Muestra los eventos cuyo nombre comienza por los caracteres insertados en la caja de texto de esta opción (ver Figura 52).

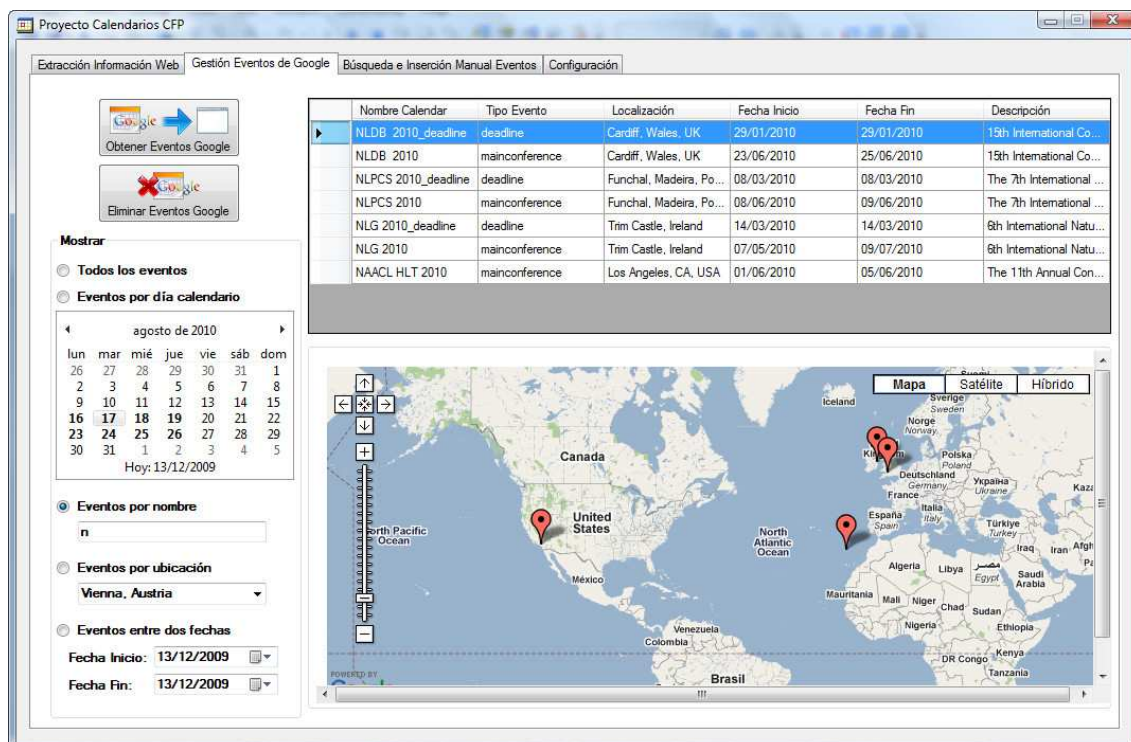


Figura 52. Mostrar los eventos que empiecen por el nombre introducido

- **Eventos por ubicación.** Muestra los eventos cuya localización coincide con la ubicación seleccionada en la lista de ubicaciones de esta opción (ver Figura 53).

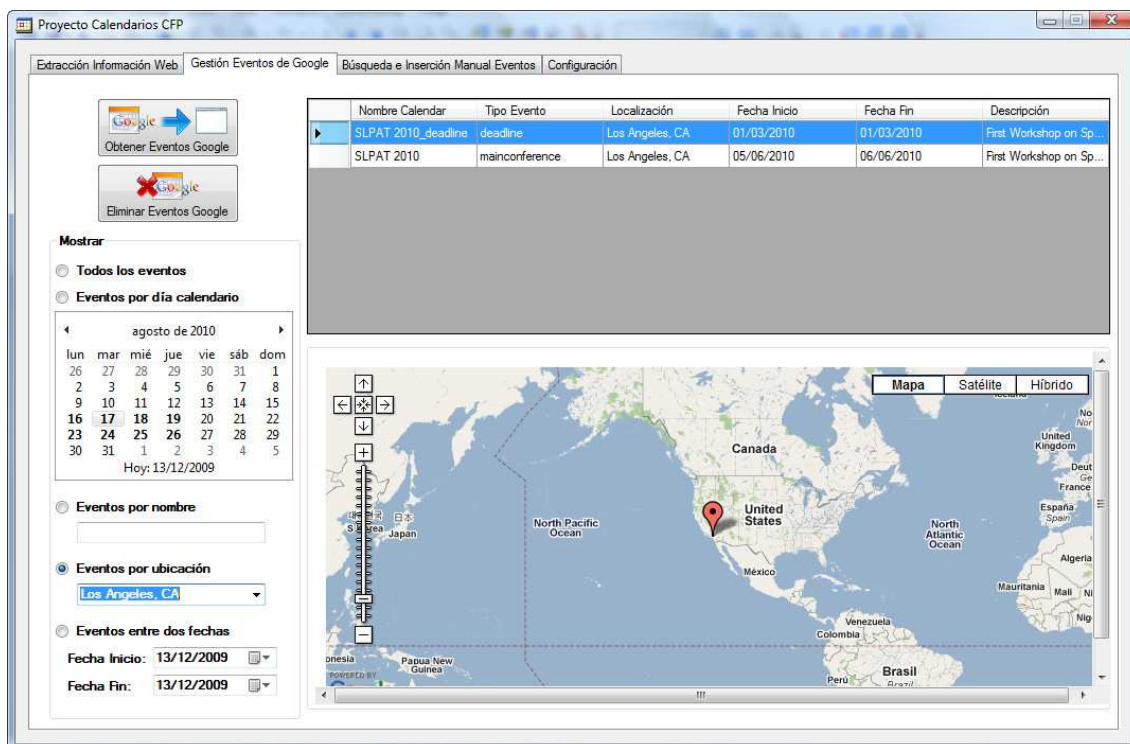


Figura 53. Mostrar los eventos filtrados por la ubicación seleccionada

- **Eventos entre dos fechas.** Muestra los eventos que tienen lugar entre las fechas especificadas en esta opción (ver Figura 54).

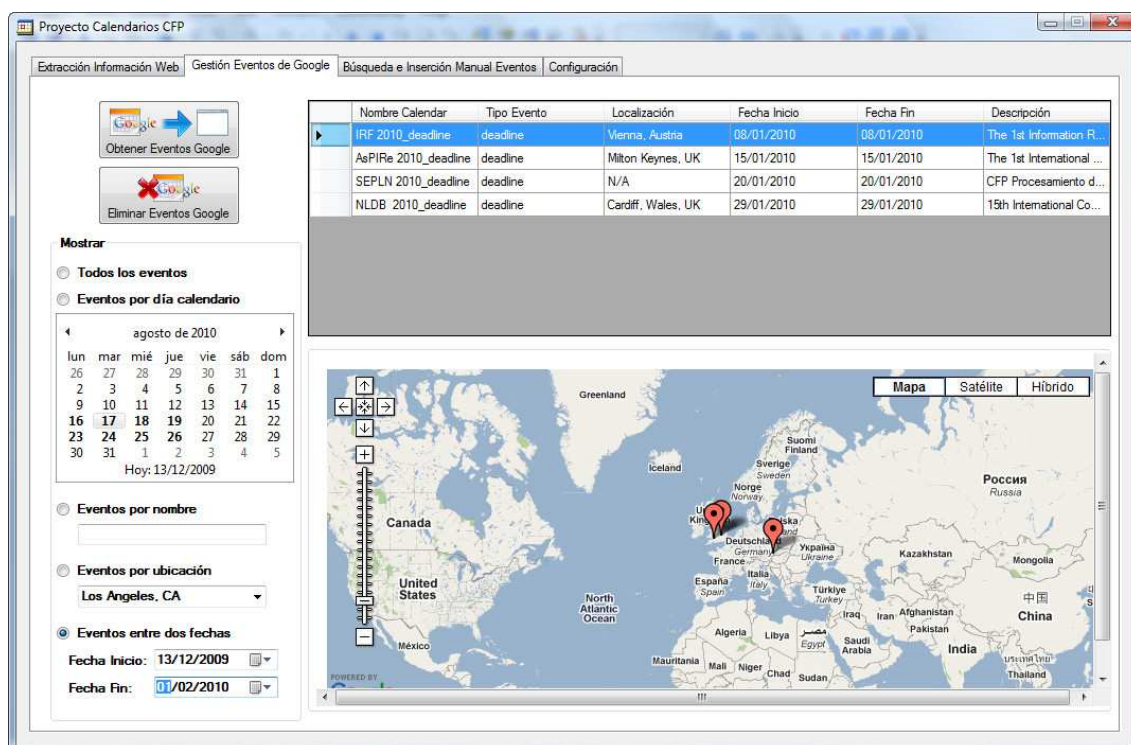


Figura 54. Mostrar los eventos que se encuentren entre el rango de fechas seleccionado

## D. Pestaña de Búsqueda e inserción manual de eventos

Por último se procede a explicar el funcionamiento de la pestaña de *Búsqueda e Inserción Manual de Eventos* (ver Figura 55).

En dicha pestaña, como se puede ver en la imagen, se pueden realizar dos operaciones:

- 1) Buscar información en páginas Web.
- 2) Insertar manualmente un nuevo evento en el calendario de Google Calendar.

### 1) Buscar información en páginas Web.

Para ello se introduce el texto a buscar en la caja de texto del buscador. Posteriormente se hace click sobre el botón *Buscar*, y automáticamente se muestran debajo las páginas encontradas. Si se hace click sobre alguna de las páginas encontradas en el buscador, se abre el explorador predeterminado del sistema cargando esa página.

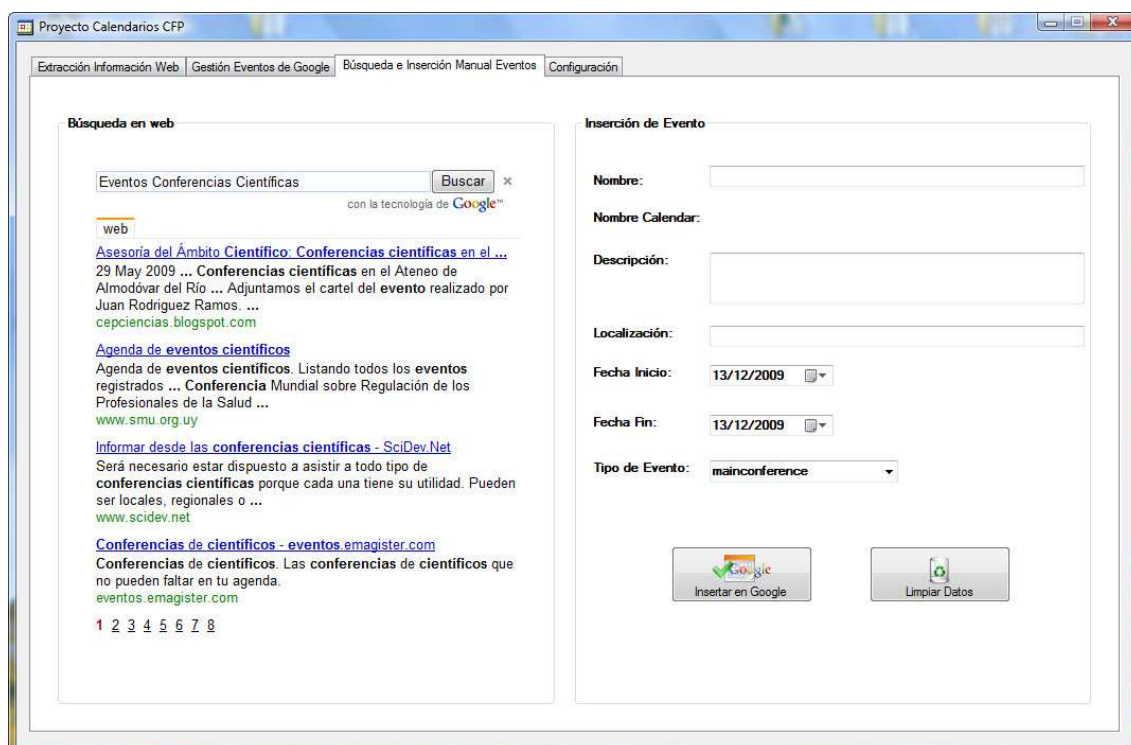


Figura 55. Pestaña de Búsqueda e Inserción Manual de Eventos

## 2) Insertar manualmente un nuevo evento en el calendario de Google Calendar.

Se rellenan los campos del formulario, se hace click sobre el botón *Insertar en Google*, y este evento se insertará dentro del calendario creado por la aplicación en Google Calendar. Para ver el evento creado dentro de la aplicación habría que ir a la pestaña *Gestión de Eventos de Google*, y hace click sobre *Obtener Eventos Google*, actualizando de esta manera los datos de la aplicación, obteniendo todos los eventos que se encuentran en el calendario de Google Calendar.

## ANEXO 2 – MANUAL DE INSTALACIÓN

Antes de comenzar con la instalación hay que tener en cuenta las siguientes consideraciones:

- La carpeta donde se instale la aplicación debe tener permisos de lectura/escritura, al igual que todas sus subcarpetas una vez acabado el proceso de instalación.

En sistemas operativos como Windows Vista, por motivos de seguridad, el sistema operativo restringe los permisos a las carpetas que se crean en la unidad C, y estos permisos hay que asignárselos manualmente, puesto que la aplicación tiene que leer y escribir en ficheros de esa ubicación para realizar determinadas operaciones. En otros sistemas operativos donde se ha probado, como Windows XP y Windows Server 2003, en principio no se ha dado este problema puesto que las carpetas no tienen restricciones de permisos para los usuarios.

- Para que la aplicación funcione correctamente es necesario que la máquina tenga instalado el Framework 2.0 de .NET. Si la máquina no lo tiene instalado, dicha operación se realizará mediante el proceso de instalación de la aplicación.

Para instalar la aplicación y el framework, si procede, hay que seguir los siguientes pasos:

1. Ejecutar el fichero ***Setup.exe***
2. Si la máquina no tiene instalado el Framework 2.0 de .NET, se seguirán los pasos que vienen a continuación. Si lo tuviera instalado, el instalador, pasará directamente a la pantalla de instalación de la aplicación. Para instalar el Framework 2.0, hacer click sobre el botón *Sí*, de la siguiente ventana (ver Figura 56):



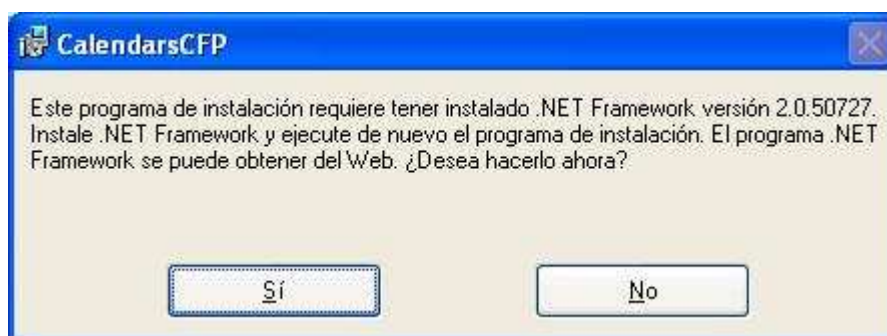


Figura 56. Comprobación de instalación de Framework 2.0 de .NET

3. Hacer click sobre el botón *Acepto*, para aceptar los términos del contrato (ver Figura 57).

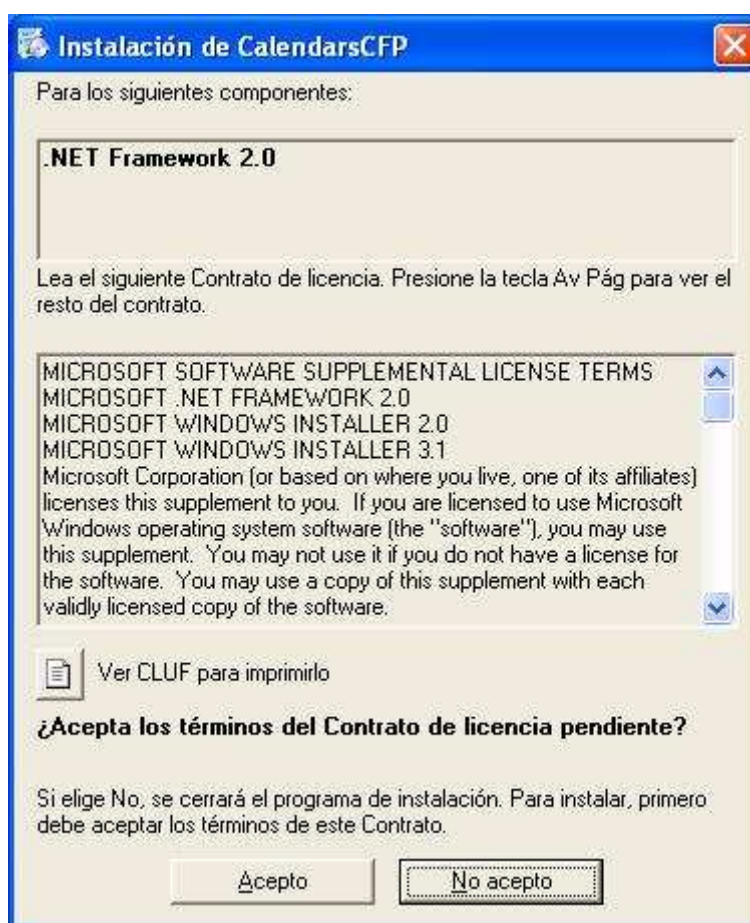
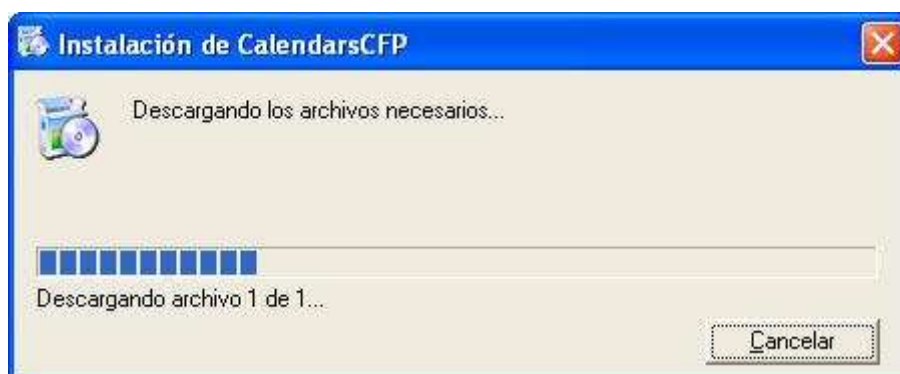


Figura 57. Aceptación de términos del contrato para la instalación de Framework 2.0 de .NET

4. A continuación se descargan de Internet los archivos necesarios (ver Figura 58).



**Figura 58.** Descarga de archivos necesarios

5. Se procede a la instalación del Framework 2.0 de .NET (ver Figura 59).



**Figura 59.** Instalación de Framework 2.0 de .NET

Los siguientes pasos se corresponden al proceso de instalación de la aplicación:

1. Hacer click sobre el botón *Siguiente* (ver Figura 60):



**Figura 60.** Instalador de la aplicación CalendarsCFP

2. Seleccionar la ruta donde se va a instalar la aplicación y la opción relativa a los usuarios de la máquina para permitirles el acceso, o no, a la aplicación. Posteriormente, hacer click sobre el botón *Siguiente* (ver Figura 61).





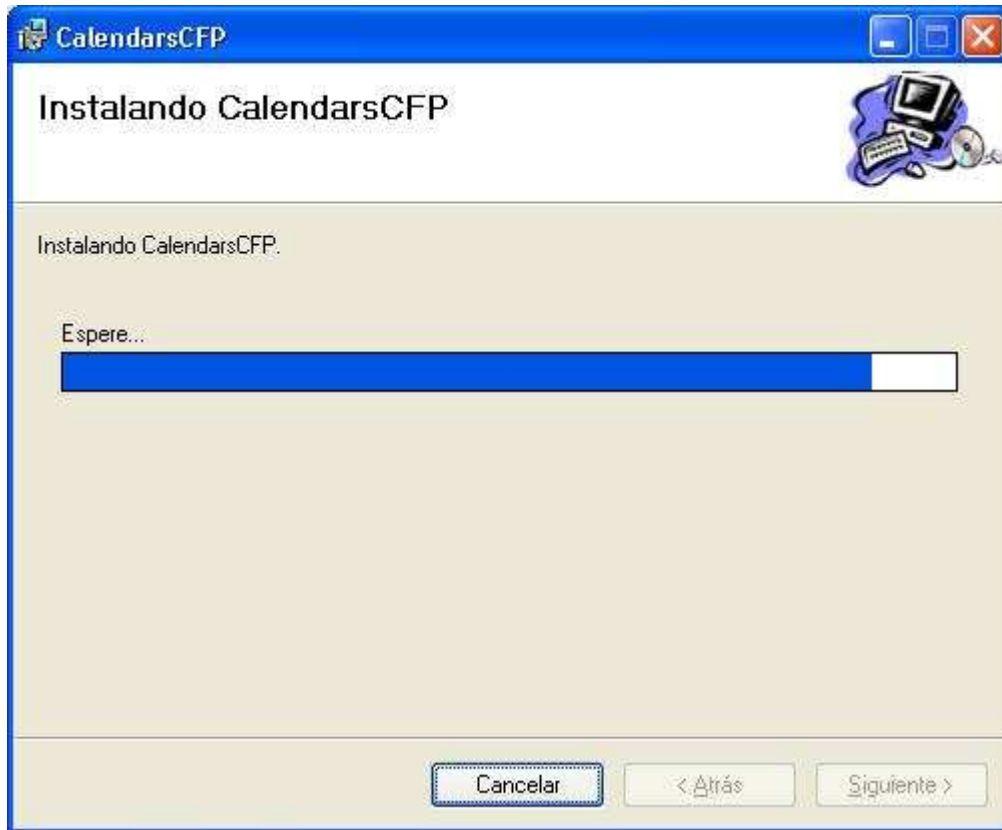
**Figura 61.** Selección de la carpeta de instalación

3. Confirmar la instalación haciendo click sobre el botón *Siguiente* (ver Figura 62).



**Figura 62.** Confirmación de la instalación de la aplicación CalendarsCFP

4. Una vez que ha terminado la barra de progreso de la instalación, hacer click sobre el botón *Siguiente* (ver Figura 63).



**Figura 63.** Instalación de la aplicación CalendarsCFP

5. La siguiente ventana indica que el proceso de instalación se ha realizado correctamente. Hacer click sobre el botón *Cerrar*, finalizando el proceso de instalación (ver Figura 64).



**Figura 64.** Instalación completada

6. Para ejecutar la aplicación hay que acceder al menú de Inicio, y dentro de éste, ir al menú CalendarsCFP y seleccionar el submenú CalendarsCFP.

## ANEXO 3 – ÍNDICE DE FIGURAS

Figura 1. Página de eventos de conferencias científicas del tipo “WikiCFP” .....	16
Figura 2. Página de eventos de conferencias científicas del tipo “ACLWiki” .....	17
Figura 3. Página de eventos de conferencias científicas del tipo “dmoz” .....	19
Figura 4. Página de eventos de conferencias científicas del tipo “aclweb” .....	20
Figura 5. Imagen de Google Calendar .....	45
Figura 6. Pantalla de Google Calendar Details para el Calendar “proyectoCalendar@gmail.com” .....	56
Figura 7. URL “Magic Cookie” para el Calendar seleccionado .....	56
Figura 8. Representación de mapa de Google .....	58
Figura 9. Representación de buscador de Google .....	58
Figura 10. Representación de feeds de Google .....	59
Figura 11. Representación de informes de Google .....	59
Figura 12. Representación de idiomas de Google .....	60
Figura 13. Representación de bibliotecas AJAX.....	60
Figura 14. Representación del API de Google para libros .....	60
Figura 15. Representación de Google Earth.....	61
Figura 16. Representación de Google Friend Connect.....	61
Figura 17. Ciclo de vida fuente .....	100
Figura 18. Diagrama de casos de uso del sistema .....	149
Figura 19. Diagrama de secuencia que describe el caso de uso CU-001 .....	157
Figura 20. Diagrama de secuencia que describe el caso de uso CU-002 .....	158
Figura 21. Diagrama de secuencia que describe el caso de uso CU-003 .....	159
Figura 22. Diagrama de secuencia que describe el caso de uso CU-004 .....	160
Figura 23. Diagrama de secuencia que describe el caso de uso CU-005 .....	161
Figura 24. Diagrama de secuencia que describe el caso de uso CU-006 .....	162
Figura 25. Diagrama de secuencia que describe el caso de uso CU-007 .....	163
Figura 26. Diagrama de secuencia que describe el caso de uso CU-008 .....	164
Figura 27. Diagrama de secuencia que describe el caso de uso CU-009 .....	165
Figura 28. Diagrama de secuencia que describe el caso de uso CU-010 .....	166
Figura 29. Diagrama de secuencia que describe el caso de uso CU-011 .....	167
Figura 30. Diagrama de secuencia que describe el caso de uso CU-012 .....	168
Figura 31. Diagrama de clases del sistema.....	169
Página 228 de 237	

Figura 32. Pestaña de extracción de información.....	186
Figura 33. Pestaña de Gestión de Eventos de Google .....	188
Figura 34. Pestaña de Búsqueda e Inserción Manual de Eventos .....	189
Figura 35. Pestaña de Configuración.....	190
Figura 36. Pestaña de datos de configuración .....	203
Figura 37. Confirmación del cambio en los datos de conexión del usuario.....	204
Figura 38. Confirmación al cambiar el nombre del calendario .....	205
Figura 39. Aviso de datos guardados correctamente.....	206
Figura 40. Pestaña de Extracción de Información Web .....	207
Figura 41. Información mostrada después de Obtener Eventos de un origen de datos .	208
Figura 42. Guardar documento XML en un fichero.....	209
Figura 43. Recuperar documento XML de un fichero.....	210
Figura 44. Pestaña de Eventos de Google .....	211
Figura 45. Pestaña de Eventos de Google después de obtener eventos.....	212
Figura 46. Aviso de confirmación de borrado de todos los eventos.....	213
Figura 47. Menú contextual asociado a un evento seleccionado en la lista .....	214
Figura 48. Confirmación de borrado del evento seleccionado .....	214
Figura 49. Formulario de edición de un evento.....	215
Figura 50. Mostrar todos los eventos.....	216
Figura 51. Mostrar los eventos para el día seleccionado en el calendario.....	217
Figura 52. Mostrar los eventos que empiecen por el nombre introducido .....	217
Figura 53. Mostrar los eventos filtrados por la ubicación seleccionada.....	218
Figura 54. Mostrar los eventos que se encuentren entre el rango de fechas seleccionado .....	219
Figura 55. Pestaña de Búsqueda e Inserción Manual de Eventos .....	220
Figura 56. Comprobación de instalación de Framework 2.0 de .NET .....	222
Figura 57. Aceptación de términos del contrato para la instalación de Framework 2.0 de .NET .....	222
Figura 58. Descarga de archivos necesarios.....	223
Figura 59. Instalación de Framework 2.0 de .NET .....	223
Figura 60. Instalador de la aplicación CalendarsCFP.....	224
Figura 61. Selección de la carpeta de instalación.....	225
Figura 62. Confirmación de la instalación de la aplicación CalendarsCFP .....	225

Figura 63. Instalación de la aplicación CalendarsCFP ..... 226

Figura 64. Instalación completada..... 227

## ANEXO 4 – ÍNDICE DE TABLAS

Tabla 1. Formato de datos de eventos de página tipo “ACLWiki” .....	18
Tabla 2. Evento mostrado en página tipo “ACLWiki” .....	18
Tabla 3. Patrón de muestra de eventos de página tipo “dmoz” .....	20
Tabla 4. Ejemplo de utilización del microformato Geo .....	33
Tabla 5. Ejemplo de representación de datos de contacto sin microformatos.....	34
Tabla 6. Ejemplo de representación de datos de contacto utilizando el microformato hCard .....	34
Tabla 7. Texto para ejemplo de microformato .....	39
Tabla 8. Código HTML sin uso de microformato .....	39
Tabla 9. Código HTML utilizando el microformato hCalendar.....	40
Tabla 10. Relación de las APIs de datos de Google con sus librerías cliente asociadas	43
Tabla 11. Relación de API y herramientas con necesidades del usuario .....	47
Tabla 12. Url de autenticación AuthSubRequest.....	52
Tabla 13. Parámetros de la URL de autenticación AuthSubRequest .....	52
Tabla 14. Url de autenticación AuthSubRequest después de que el usuario se haya autenticado.....	53
Tabla 15. URL de petición POST para autenticación ClientLogin .....	53
Tabla 16. Parámetros del cuerpo de la petición POST para autenticación ClientLogin	54
Tabla 17. Cabecera de autorización.....	54
Tabla 18. URL “Magic Cookie” .....	56
Tabla 19. Url que contiene la clave del API de Google .....	63
Tabla 20. Ejemplo de utilización del método google.load .....	63
Tabla 21. Ejemplo de selección de idioma del API de Google Maps .....	64
Tabla 22. Ejemplo de comienzo de un documento XHTML .....	66
Tabla 23. Definición de identificadores de tipos de requisitos.....	82
Tabla 24. Especificación de requisitos de usuario.....	83
Tabla 25. Especificación de requisitos del sistema .....	87
Tabla 26. Descripción de las prueba del sistema.....	90
Tabla 27. Descripción de la prueba PR-001 .....	91
Tabla 28. Descripción de la prueba PR-002 .....	91
Tabla 29. Descripción de la prueba PR-003 .....	92
Tabla 30. Descripción de la prueba PR-004 .....	92
Página 231 de 237	

Tabla 31. Descripción de la prueba PR-005 .....	93
Tabla 32. Descripción de la prueba PR-006 .....	93
Tabla 33. Descripción de la prueba PR-007 .....	94
Tabla 34. Descripción de la prueba PR-008 .....	94
Tabla 35. Descripción de la prueba PR-009 .....	95
Tabla 36. Descripción de la prueba PR-010 .....	95
Tabla 37. Descripción de la prueba PR-011 .....	96
Tabla 38. Descripción de la prueba PR-012 .....	96
Tabla 39. Descripción de la prueba PR-013 .....	97
Tabla 40. Descripción de la prueba PR-014 .....	97
Tabla 41. Descripción de la prueba PR-015 .....	98
Tabla 42. Descripción de la prueba PR-016 .....	98
Tabla 43. Descripción del requisito RCP-001 .....	104
Tabla 44. Descripción del requisito RCP-002 .....	105
Tabla 45. Descripción del requisito RCP-003 .....	105
Tabla 46. Descripción del requisito RCP-004 .....	106
Tabla 47. Descripción del requisito RCP-005 .....	106
Tabla 48. Descripción del requisito RCP-006 .....	107
Tabla 49. Descripción del requisito RRS-001 .....	107
Tabla 50. Descripción del requisito RRS-002 .....	108
Tabla 51. Descripción del requisito RFN-001 .....	108
Tabla 52. Descripción del requisito RFN-002 .....	109
Tabla 53. Descripción del requisito RFN-003 .....	109
Tabla 54. Descripción del requisito RFN-004 .....	110
Tabla 55. Descripción del requisito RFN-005 .....	110
Tabla 56. Descripción del requisito RFN-006 .....	111
Tabla 57. Descripción del requisito RFN-007 .....	111
Tabla 58. Descripción del requisito RFN-008 .....	112
Tabla 59. Descripción del requisito RFN-009 .....	112
Tabla 60. Descripción del requisito RFN-010 .....	113
Tabla 61. Descripción del requisito RFN-011 .....	113
Tabla 62. Descripción del requisito RFN-012 .....	114
Tabla 63. Descripción del requisito RFN-013 .....	114



Tabla 64. Descripción del requisito RFN-014.....	115
Tabla 65. Descripción del requisito RFN-015.....	115
Tabla 66. Descripción del requisito RFN-016.....	115
Tabla 67. Descripción del requisito RFN-017.....	116
Tabla 68. Descripción del requisito RFN-018.....	116
Tabla 69. Descripción del requisito RRD-001 .....	117
Tabla 70. Descripción del requisito RRD-002 .....	117
Tabla 71. Descripción del requisito RIN-001.....	118
Tabla 72. Descripción del requisito RIN-002.....	118
Tabla 73. Descripción del requisito RIN-003.....	119
Tabla 74. Descripción del requisito RIN-004.....	119
Tabla 75. Descripción del requisito RIN-005.....	120
Tabla 76. Descripción del requisito RIN-006.....	120
Tabla 77. Descripción del requisito RIN-007.....	120
Tabla 78. Descripción del requisito RIN-008.....	121
Tabla 79. Descripción del requisito RIN-009.....	121
Tabla 80. Descripción del requisito RIN-010.....	122
Tabla 81. Descripción del requisito RIN-011.....	122
Tabla 82. Descripción del requisito RIN-012.....	123
Tabla 83. Descripción del requisito RIN-013.....	123
Tabla 84. Descripción del requisito RIN-014.....	124
Tabla 85. Descripción del requisito RIN-015.....	124
Tabla 86. Descripción del requisito RIN-016.....	125
Tabla 87. Descripción del requisito RIN-017.....	125
Tabla 88. Descripción del requisito RIN-018.....	126
Tabla 89. Descripción del requisito RIN-019.....	126
Tabla 90. Descripción del requisito RIN-020.....	127
Tabla 91. Descripción del requisito RIN-021.....	127
Tabla 92. Descripción del requisito RIN-022.....	128
Tabla 93. Descripción del requisito ROP-001 .....	128
Tabla 94. Descripción del requisito ROP-002.....	129
Tabla 95. Descripción del requisito ROP-003.....	129
Tabla 96. Descripción del requisito ROP-004.....	129

Tabla 97. Descripción del requisito ROP-005.....	130
Tabla 98. Descripción del requisito ROP-006.....	130
Tabla 99. Descripción del requisito ROP-007.....	131
Tabla 100. Descripción del requisito ROP-008.....	131
Tabla 101. Descripción del requisito ROP-009.....	132
Tabla 102. Descripción del requisito ROP-010.....	132
Tabla 103. Descripción del requisito ROP-011.....	133
Tabla 104. Descripción del requisito ROP-012.....	133
Tabla 105. Descripción del requisito ROP-013.....	134
Tabla 106. Descripción del requisito ROP-014.....	134
Tabla 107. Descripción del requisito ROP-015.....	135
Tabla 108. Descripción del requisito ROP-016.....	135
Tabla 109. Descripción del requisito ROP-017.....	136
Tabla 110. Descripción del requisito ROP-018.....	136
Tabla 111. Descripción del requisito ROP-019.....	137
Tabla 112. Descripción del requisito ROP-020.....	137
Tabla 113. Descripción del requisito ROP-021.....	138
Tabla 114. Descripción del requisito ROP-022.....	138
Tabla 115. Descripción del requisito ROP-023.....	139
Tabla 116. Descripción del requisito ROP-024.....	139
Tabla 117. Descripción del requisito ROP-025.....	140
Tabla 118. Descripción del requisito ROP-026.....	140
Tabla 119. Descripción del requisito RRC-001 .....	141
Tabla 120. Descripción del requisito RCM-001.....	141
Tabla 121. Descripción del requisito RCM-002.....	142
Tabla 122. Descripción del requisito RCM-003.....	142
Tabla 123. Descripción del requisito RCM-004.....	143
Tabla 124. Descripción del requisito RCM-005.....	143
Tabla 125. Descripción del requisito RCM-006.....	144
Tabla 126. Descripción del requisito RDC-001 .....	144
Tabla 127. Descripción del requisito RDC-002 .....	145
Tabla 128. Descripción del requisito RDC-003 .....	145
Tabla 129. Descripción del requisito RDC-004 .....	145

Tabla 130. Descripción del requisito RDC-005 .....	146
Tabla 131. Descripción del requisito RDC-006 .....	146
Tabla 132. Descripción del requisito RSG-001 .....	146
Tabla 133. Descripción del requisito RSG-002.....	147
Tabla 134. Tabla de identificación de casos de uso del sistema.....	148
Tabla 135. Descripción del caso de uso CU-001.....	150
Tabla 136. Descripción del caso de uso CU-002.....	151
Tabla 137. Descripción del caso de uso CU-003.....	151
Tabla 138. Descripción del caso de uso CU-004.....	152
Tabla 139. Descripción del caso de uso CU-005.....	152
Tabla 140. Descripción del caso de uso CU-006.....	153
Tabla 141. Descripción del caso de uso CU-007.....	153
Tabla 142. Descripción del caso de uso CU-008.....	154
Tabla 143. Descripción del caso de uso CU-009.....	154
Tabla 144. Descripción del caso de uso CU-010.....	155
Tabla 145. Descripción del caso de uso CU-011.....	155
Tabla 146. Descripción del caso de uso CU-012.....	156
Tabla 147. Descripción de las clases del sistema .....	170
Tabla 148. Descripción de las propiedades de la clase Definiciones .....	171
Tabla 149. Descripción de las propiedades que contiene la estructura Evento.....	172
Tabla 150. Descripción de los valores de TIPOSEVENTOS.....	172
Tabla 151. Descripción de los métodos de la clase Definiciones.....	172
Tabla 152. Descripción de las propiedades de la clase ExtraccionInformacion .....	173
Tabla 153. Descripción de los valores de TIPOSPAGINAS .....	173
Tabla 154. Descripción de los métodos de la clase ExtraccionInformacion .....	174
Tabla 155. Descripción de las propiedades públicas de la clase Calendar.....	174
Tabla 156. Descripción de los métodos de la clase Calendar.....	176
Tabla 157. Descripción de los métodos de la clase Configuracion .....	177
Tabla 158. Descripción de las propiedades de la clase FrmEvento.....	177
Tabla 159. Descripción de los métodos de la clase FrmEvento .....	178
Tabla 160. Descripción de los métodos de la clase FrmPrincipal .....	178
Tabla 161. Descripción de los métodos de la clase gMaps .....	179
Tabla 162. Descripción de los métodos de la clase TrazaLog.....	179

Tabla 163. Descripción de los métodos de la clase TripleDES .....	180
Tabla 164. Esquema de validación (XSD) aplicado al documento XML de eventos ..	184
Tabla 165. Resultados de las pruebas de validación del sistema.....	191

